

Appendix A

```

; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-045-583-52

Query Match      84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
Db 1 MDYQVSSPIYDIDITISPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFSGIFII 120
Db 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFSGIFII 120
QY 121 LTTIDRYLVAVHAFALKARTVFGVVTSTVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
Db 121 LTTIDRYLVAVHAFALKARTVFGVVTSTVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
QY 181 HPFY 184
Db 181 HPFY 184

RESULT 9
US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptahelical Receptor Superfamily and Uses
; therefor
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/534,185
; FILING DATE: 24-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,583
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-534-185-52

```

```

Query Match      84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
Db 1 MDYQVSSPIYDIDITISPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFSGIFII 120
Db 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFSGIFII 120
QY 121 LTTIDRYLVAVHAFALKARTVFGVVTSTVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
Db 121 LTTIDRYLVAVHAFALKARTVFGVVTSTVITWVAVFASLPGLIIFTRSQKGLHYTCSS 180
QY 181 HPFY 184
Db 181 HPFY 184

RESULT 10
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; CHEMOKINE RECEPTOR HDGRL0 (AS AMENDED)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 433
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-343D-2

Query Match      84.0%; Score 943; DB 3; Length 352;
Best Local Similarity 97.8%; Pred. No. 2.9e-84;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
Db 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFSGIFII 120
Db 61 LKSMTDIYLLNLALISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFSGIFII 120

```

| | | | |
|----|-----|---|-----|
| Qy | 121 | LITIDRYLVAVHAFALKARVTEGVVTSIIIVWVAFASIPGIIFTRSQEGILHYTCSS | 180 |
| | | | |
| | | | |
| Db | 121 | LITIDRYLVAVHAFALKARVTEGVVTSIIIVWVAFASIPGIIFTRSQEGILHYTCSS | 180 |
| | | | |
| | | | |
| Qy | 181 | HPFY 184 | |
| | | | |
| Db | 181 | HPFY 184 | |
| | | | |

```

RESULT 11
US-09-517-605-5
; Sequence 5, Application US/09517605
; Patent NO. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517.605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-5

```

RESULT 12
 US-08-724-984A-2
 ; Sequence 2, Application US/08724984A
 ; Patent No. 6388055
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon
 ;
 ; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
 ;
 ; TITLE OF INVENTION: CKR5 Receptor
 ;
 ; NUMBER OF SEQUENCES: 5
 ;
 ; CORRESPONDENCE ADDRESS:
 ;
 ; ADDRESSEE: SmithKline Beecham Corporation
 ;
 ; STREET: 709 Swedeland Road, P.O. Box 1539
 ;
 ; CITY: King of Prussia
 ;
 ; STATE: PA
 ;
 ; COUNTRY: USA
 ;
 ; ZIP: 19406-0939
 ;
 ; COMPUTER READABLE FORM:
 ;
 ; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
 ;
 ; COMPUTER: IBM 486
 ;
 ; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

```

SOFTWARE: MICROSOFT WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,984A
FILING DATE: October 3, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: William T. Han
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270 5024
TELEFAX: 610 270 5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 354
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-724-984A-2

Query Match          69.2%; Score 776; DB 4; Length 354;
Best Local Similarity 79.0%; Pred. No. 5.3e-68;
Matches 147; Conservative 17; Mismatches 20; Indels 2; Gaps 1;

QY      1 MDYQVSSP--IYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFVGNMLVILLINC 58
Db      1 MDFQSGVPTIYDIDYGSAPCQKINVKQIAAQLPPLYSLVIFGFAGNMVFLILISC 60

QY      59 KRLKSWTDIYLLNLAISDLFELLIVPFWAHYAAQWDFGNTMCLLTGLGYFGFSGIEFF 118
Db      61 KKIKSWTDIYLLNLAISDLFLTLFPWARYAANEWIFGNIMCKVFTGVYHIGYFGGIEF 120

QY      119 IILLIDRYLAVYHVFALKARTYTFGWTSVITWVAVFASLPGLIIFTRSQEGLHYTC 178
Db      121 IILLIDRYLAVYHVFALKYTTVFCGVITSVTWVAVFASLPGLIIFTRSQEGLHYTC 180

QY      179 SSHPY 184
Db      181 SPHPH 186

```

RESULT 13
US-08-466-343D-9
Sequence 9, Application US/08466343D
Patent No. 6025154
GENERAL INFORMATION:
APPLICANT: LI, Yi
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGRR10 (AS AMENDED)
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSES: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,343D
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
TELECOMMUNICATION INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:39:24 ; Search time 74.6348 Seconds
(without alignments)
5149.005 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2828
Sequence: 1 GAATTCCTCCCAACAGAGCCA.....AGTAGATTAGATCCGATTC 1442

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09938703/runat_03062003_161406_22796/app.query.fasta_1.1870
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptn -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.qcgn_1_162.runat_03062003_161406_22796 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq_emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq_emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq_emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq_emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq_emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq_emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq_emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq_emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq_emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq_emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq_emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq_emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq_emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq_emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq_emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq_emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq_emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq_emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq_emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq_emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq_emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq_emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq_emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------|---------------------|
| 1 | 1768.5 | 67.3 | 371 | 19 AAW23834 | Human CC chemokine |
| 2 | 1745.5 | 66.4 | 352 | 18 AAW27407 | Human CCR5. Homo |
| 3 | 1745.5 | 66.4 | 352 | 18 AAW27123 | Human chemokine re |
| 4 | 1745.5 | 66.4 | 352 | 19 AAW23835 | Human CC chemokine |
| 5 | 1745.5 | 66.4 | 352 | 20 AAW88232 | HIV-1 co-receptor |
| 6 | 1745.5 | 66.4 | 352 | 22 AAG80111 | Human CCR5 protein |
| 7 | 1745.5 | 66.4 | 352 | 22 AAB82948 | Human HIV-1 co-rec |
| 8 | 1745.5 | 66.4 | 352 | 22 AAB83354 | Human CCR5 protein |
| 9 | 1745.5 | 66.4 | 352 | 22 AAE04321 | Human chemokine re |
| 10 | 1745.5 | 66.4 | 352 | 23 ABB08343 | Human chemokine (C |
| 11 | 1745.5 | 66.4 | 352 | 23 AAM52828 | Human CC chemokine |
| 12 | 1745.5 | 66.4 | 439 | 20 AAY41280 | Fusion protein con |
| 13 | 1739.5 | 66.2 | 352 | 22 ABB56342 | Non-endogenous hum |
| 14 | 1739.5 | 66.2 | 352 | 23 AAM52829 | Human CCR5 Gln 55 |
| 15 | 1737.5 | 66.1 | 352 | 22 AAE07039 | Human G-protein ch |
| 16 | 1737.5 | 66.1 | 352 | 22 AAE07048 | Human G-protein ch |
| 17 | 1737.5 | 66.1 | 352 | 22 AAB46858 | Human HDGRI10 prot |
| 18 | 1737.5 | 66.1 | 352 | 23 AAG97152 | Human G-protein ch |
| 19 | 1730.5 | 65.8 | 352 | 18 AAW07602 | Human G-protein ch |
| 20 | 1730.5 | 65.8 | 352 | 21 AAY80128 | Human G-protein ch |
| 21 | 1730.5 | 65.8 | 352 | 22 AAE07037 | Human G-protein ch |
| 22 | 1730.5 | 65.8 | 352 | 22 AAE07046 | Human G-protein ch |
| 23 | 1730.5 | 65.8 | 352 | 23 AAG97150 | Human G-protein ch |
| 24 | 1718.5 | 65.4 | 352 | 22 AAG75089 | Amino acid sequenc |
| 25 | 1711.5 | 65.1 | 352 | 18 AAW27125 | Macaque chemokine |
| 26 | 1670 | 63.5 | 332 | 18 AAW26766 | Human chemokine re |
| 27 | 1463.5 | 55.7 | 354 | 19 AAW54037 | Mouse CC-CKR5 prot |
| 28 | 1332 | 50.7 | 360 | 22 AAG07613 | Human CCR2-641 pol |
| 29 | 1331 | 50.6 | 360 | 16 AAR79166 | Human monocyte che |
| 30 | 1331 | 50.6 | 360 | 18 AAW35833 | Human monocyte che |
| 31 | 1331 | 50.6 | 360 | 22 AAG80108 | Human CCR2b protei |
| 32 | 1331 | 50.6 | 360 | 22 AAG07614 | Human wild-type CC |
| 33 | 1325 | 50.4 | 360 | 22 ABB56340 | Non-endogenous hum |
| 34 | 1191 | 45.3 | 374 | 16 AAR79165 | Human monocyte che |
| 35 | 1191 | 45.3 | 374 | 22 AAG80107 | Human CCR2a protei |
| 36 | 1122 | 42.7 | 215 | 18 AAW27408 | Inactive human CCR |
| 37 | 1122 | 42.7 | 215 | 20 AAW88238 | HIV-1 co-receptor |
| 38 | 1105 | 42.0 | 329 | 22 AAB46859 | Human MCP-1 recept |
| 39 | 1003.5 | 38.2 | 355 | 15 AAR52749 | C-C chemokine rece |
| 40 | 1003.5 | 38.2 | 355 | 18 AAW26588 | Human MIP-1 alpha/R |
| 41 | 1003.5 | 38.2 | 355 | 18 AAW25751 | Human MIP-1 alpha/R |
| 42 | 1003.5 | 38.2 | 355 | 21 AAB20571 | Human CC-chemokine |
| 43 | 965 | 36.7 | 355 | 18 AAW29179 | Rat CC chemokine r |
| 44 | 958 | 36.5 | 184 | 18 AAW27406 | Inactive human CCR |
| 45 | 916 | 34.9 | 355 | 19 AAW51744 | Human C-C chemokin |

ALIGNMENTS

RESULT 1
AAW23834
ID AAW23834 standard; Protein; 371 AA.
XX AC AAW23834;
XX AC AAW23834;
XX AC AAW23834;
DT 08-JUN-1998 (first entry)
XX DE Human CC chemokine receptor 5 (CCR5) A127V variant.
XX DE CC chemokine receptor 5; CCR5; G-protein coupled receptor;
XX KW human immunodeficiency virus; HIV; CCR5; AIDS; therapy;
XX KW transgenic animal.
XX OS Homo sapiens.
XX EH Key
XX FT Location/Qualifiers
FT Domain 48..74
FT /label= I

FT /note= "transmembrane domain"
 FT 123..145
 FT /label= III
 FT /note= "transmembrane domain"
 FT 162..190
 FT /label= IV
 FT /note= "transmembrane domain"
 FT 213..238
 FT /label= V
 FT /note= "transmembrane domain"
 FT 257..277
 FT /label= VI
 FT /note= "transmembrane domain"
 FT 296..319
 FT /label= VII
 FT /note= "transmembrane domain"

PN W09745543-A2.

PD 04-DEC-1997.

PF 28-MAY-1997; 97WO-US09586.

PR 28-MAY-1996; 96US-0018508.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Alkhatib G, Berger EA, Broder CC, Combadiere C;

PI Feng Y, Kennedy PE, Murphy PM;

DR WPI; 1998-032650/03.

DR N-PSDB; AAT76919.

PT CC chemokine receptor 5 polypeptide - used to inhibit membrane
 PT fusion between HIV and a target cell

PS Example 1; Fig 1b; 70pp; English.

CC This protein sequence comprises an Alal27Leu variant of a novel
 CC human macrophage-selective CC chemokine receptor (see also
 CC AAW23835) that has been designated CCR5. The sequence was deduced
 CC from an isolated cDNA clone (see AAT76919). The conservative
 CC variation should not affect the activity of CCR5. The susceptibility
 CC of human macrophages to HIV infection depends on cell surface
 CC expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane
 CC superfamily of G-protein coupled cell surface molecules. It plays
 CC an essential role in the membrane fusion step of infection by some
 CC HIV isolates. The establishment of stable, non-human cell lines
 CC and transgenic mammals having cells that coexpress human CD4 and
 CC CCR5 provides valuable tools for research of HIV infection.
 CC Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding
 CC agents capable of blocking membrane fusion between HIV and target
 CC cells represent potential anti-HIV therapeutics for macrophage
 CC tropic strains of HIV.

XX SQ Sequence 371 AA;

Alignment Scores:
 Pred. No.: 8, 91e-190 Length: 371
 Score: 1768.50 Matches: 343
 Percent Similarity: 96.36%
 Best Local Similarity: 96.08%
 Query Match: 67.29%
 DB: 19 Gaps: 1

US-09-938-703-3 (1-1442) x AAW23834 (1-371)

QY 228 GGTGGACACAGATGGATATCAAGTGTCAGAGTCCATCTATGACATCAATATTATACA 287
 DB 16 GlyTriaAsnLysMetAspTyrGlnValSerProIleTyrAspIleAsnTyrTyrThr 35
 QY 288 TCGAGGCGCTGCGAAATCAATGTGAGCAATATCGACCGCGCTCTCCGCCGCTC 347

Db 36 SerGluProCysGlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeu 55
 QY 348 TACTCACTGGTGTTCATCTTTGGTTTGGGCAACATCTGCTGTCATCTCTCATCTGATA 407
 Db 56 TyrSerLeuValPheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeu 75
 QY 408 AACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTGAC 467
 Db 76 AsnCysLysArgLeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAsp 95
 QY 468 CTGTTTTCCTCTTACTCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
 Db 96 LeuPhePheLeuLeuThrValProPheThrPalaHisTyrLeuAlaIleGlnTrpAspPhe 115
 QY 528 GGAAATACAAATGTGTCAACTCTGTGACAGGCTCTATTTTATAGCTTCTCTCTGGAATC 587
 Db 116 GlyAsnThrMetCysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIle 135
 QY 588 TTCTTCATCATCTCTCTGACATCTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 647
 Db 136 PhePheIleIleLeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAla 155
 QY 648 TTAAAAGCCAGGACGGTCACTCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 707
 Db 156 LeuLysAlaArgThrValThrPheGlyValValThrSerValIleThrTrpValValAla 175
 QY 708 GTGTTTGGCTCTCTCCAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCAATAC 767
 Db 176 ValPheAlaSerLeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyr 195
 QY 768 ACCTGACGCTCTCATTTTCCATAC-----ATT 794
 Db 196 ThrCysSerSerHisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu 215
 QY 795 AAAGATAGTATCTCTTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 854
 Db 216 -LysIleValIleLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGln 235
 QY 855 AATCTTAAACTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 914
 Db 235 YlleLeuLysThrLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgIle 255
 QY 915 TATCTTCACCATCATGATGTTTATTTTCTCTGCTGGCTCCCTACACATTTGCTTCT 974
 Db 255 uilePheThrIleMetIleValThrPheLeuPheThrPalaProTyrAsnIleValLeu 275
 QY 975 CCTGAACACTTCCAGGAATCTTTGGCTGGAATTCGAGTACGCTCTAACAGGTGGA 1034
 Db 275 uLeuAsnThrPheGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAs 295
 QY 1035 CCAAGCTATGACAGGTGACAGACTCTTGGGATGACGACTGCTGCATCAACCCATCAT 1094
 Db 295 pGlnAlaMetGlnValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIle 315
 QY 1095 CTATGCTCTTGTGGGAGAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACAT 1154
 Db 315 eTyrAlaPheValGlyGluLysPheArgAsnTyrLeuValPhePheGlnLysHisIle 335
 QY 1155 TGCCAAACGGTCTGCAAAATGCTGTTCTATTTCAGAGAGGCTCCGAGGAGGCAAG 1214
 Db 335 eAlaLysArgPheCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlase 355
 QY 1215 CTCAGTTTACCCGATCCATCCATGGGAGGAGGAGAAATATCTGTGGGCTTG 1263
 Db 355 rSerValTyrThrArgSerThrGlyGluGlnGluIleSerValGlyIleu 371

RESULT 2

AAW27407

ID AAW27407 standard; Protein; 352 AA.

XX AC AAW27407;

XX AC AAW27407;

DT 14-APR-1998 (first entry)

XX DE Human CCR5.
 XX DE
 KW Human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
 KW diagnosis; treatment; prevention;
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder.
 XX OS
 XX OS Homo sapiens.
 XX PN W09732019-A2.
 XX PD 04-SEP-1997.
 XX PF 28-FEB-1997; 97WO-BE00023.
 XX PR 06-AUG-1996; 96EP-0870102.
 XX PR 01-MAR-1996; 96EP-0870021.
 XX PR (EURO-) EUROSREEN SA.
 XX PI Libert F, Parmentier M, Samson M, Vassart G;
 XX WPI; 1997-479829/44.
 XX DR N-PSDB; AAT90117.
 XX PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 PT useful to diagnose, prevent and/or treat inflammatory disorders,
 PT autoimmune disease and viral infection
 XX PS Claim 4; Fig 1b-c; 94pp; English.
 XX CC The present sequence is human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),
 CC MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis,
 CC glomerulonephritis, asthma, idiopathic pulmonary fibrosis and
 CC psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 XX SQ Sequence 352 AA;
 Alignment Scores:
 Pred. No.: 3 38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 18 Gaps: 1
 US-09-938-703-3 (1-1442) x AAW27407 (1-352)
 QY 240 ATGATATCAAGTGTCAAGTCCATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAATATCAATGAGCAATCGCAGCGCCCTCCCTGCTCCGCTCACTCACTGGTG 359
 DB 21 GlnLysIleAsnValIysGlnIleAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 40
 QY 360 TTCATCTTTGGTTTGGGCAACATGCTGTCATCTCACTCACTCACTGATAACTGCAAAAGG 419
 DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCCCTT 479
 DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80

QY 480 CTTACTGTCCCTTCTGGCTCACTATGCTGCCGCCACCTGGGACTTTTGGAAATACAATG 539
 DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCACACTCTTACAGGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATC 599
 DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCTGACAAATCGATAGTACCTGGCTGCTGCATGCTGCTGTTTAAAGCCAGG 659
 DB 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTGGGCTGCTCACAAGTGTGATCACTGGGTGGTGGTGGTGGTGGTGGT 719
 DB 141 ThrValThrPheGlyValValThrSerValIleIleThrIleValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTACCATGCTCAAAAGAAGTCTTCAITACACCTGACGCTCT 779
 DB 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnLysGlnLysIleTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAI 806
 DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle 200
 QY 807 CTGGGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyLeuLysTh 220
 QY 867 TCTGCTTGGTGTGCGAAATAGAGAAGAGGACAGGCTGCTGAGGCTTATCTTCCACAT 926
 DB 220 rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrIle 240
 QY 927 CATGATTTGTTTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 DB 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTGCTGAATAATTGCAGTAGCTCTAACAGTGTGACCAAGCTATGCA 1046
 DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGl 280
 QY 1047 GGTGACAGACACTTGGGATGACGCACTGCTGCTGATCAACCCCATCATCTATGCTTTGT 1106
 DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGACTTCAGAACTACCTCTTAGTCTTCTCCAAAAGCACATTCGCAACGCTT 1166
 DB 300 lGlyLysPheArgAsnTyrLeuLeuValPheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTTCTTATTTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 DB 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGGAAATATCTGGGCTTG 1263
 DB 340 rArgSerThrGlyGlnGlnIleSerValGlyLeu 352
 RESULT 3
 AAW27123
 ID AAW27123 standard; Protein; 352 AA.
 XX AC AAW27123;
 XX AC AAW27123;
 DT 14-DEC-1997 (first entry)
 XX DE Human chemokine receptor 88C.
 XX KW Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
 KW tumour; asthma; viral infection; AIDS; inflammation;
 KW autoimmune disease; therapy; diagnosis; leukocyte trafficking;
 KW G protein coupled receptor; ligand; modulator; antibody; human.
 XX OS Homo sapiens.

Wed Jun 11 09:06:08 2003

XX Key Location/Qualifiers
PH Domain 1..32
FT /label= Extracellular_domain
FT 56..67
FT /label= Intracellular_domain
FT 89..112
FT /label= Extracellular_domain
FT 125..145
FT /label= Intracellular_domain
FT 166..191
FT /label= Extracellular_domain
FT 213..235
FT /label= Intracellular_domain
FT 259..280
FT /label= Extracellular_domain
FT 301..352
FT /label= Intracellular_domain

WO9722698-A2.

26-JUN-1997.

20-DEC-1996; 96WO-US20759.

07-JUN-1996; 96US-0661393.

20-DEC-1995; 95US-0575967.

(ICOS-) ICOS CORP.

Gray PW, Raport CJ, Schweickart VL;

WPI; 1997-341689/31.

N-PSDB; AAT85161.

New nucleic acid encoding chemokine receptors 88-2B and 88C - used to modulate leukocyte trafficking, e.g. for treatment of inflammation, tumours, viral infections, autoimmune diseases, etc.

Claim 16; Page 47-48; 65pp; English.

This polypeptide sequence comprises novel human chemokine receptor 88C, a G protein coupled receptor that is involved in leukocyte trafficking. Its amino sequence was deduced from a cDNA clone (AAT85161) isolated from a macrophage library. It shows 62% identity to CCRX1. Chemokine receptor 88-2B (see AAW27124) has also been identified. 88C and 88-2B receptors and their polypeptide fragments can be produced in transformed host cells. The receptors, peptides comprising one or more of the extracellular or intracellular domains, and anti-receptor antibodies can be used to modulate receptor activities, particularly ligand and G protein binding, and are potentially useful in the treatment of atherosclerosis, rheumatoid arthritis, tumours, asthma, viral infection, AIDS, inflammatory conditions, pathological immune response, abnormal haematopoietic processes etc.

Sequence 352 AA;

Alignment Scores:
Pred. No.: 3,38e-187 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 18 Gaps: 1

US-09-938-703-3 (1-1442) x AAW27123 (1-352)

QY 240 ATGGATTATCAAGTCAAGTCCCAATCTATGACATCAATATATATACATCGAGCCCTGCC 299

DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20

QY 300 CAAATAATCAATGTGAAGCAATCGACCGCGCTCTCTCCCTCTACTACTCTACTCTG 359

Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATTAACCTCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTTACTGTCCTCTTGGGCTACTATGCTGCCCGCCAGTGGGACCTTTGGAAATACATG 539
Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACGGGCTCTATTTATAGGCTTCTTCTCTGGAATCTTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGACAAATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTACCTTTGGGCTGACAGTGTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValAlaValPheAlaSer 160
QY 720 CTCGCCAGGAATCATCTTACCATGATCTCAAAAGAGGCTTCTTATTTACCTGCTGCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSer 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeuLysIleVal 200
QY 807 CTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
QY 867 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThr 240
QY 927 CATGATGTTTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheThrPheAlaProTyrAsnIleValLeuLeuAsnThr 260
QY 987 CCAGGAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMet 280
QY 1047 GGTGACAGACACTTGTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPhe 300
QY 1107 CGGGAGAGAGTTCAGAACTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1166
Db 300 lGlyLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArg 320
QY 1167 CTGCAATGCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1226
Db 320 ecLysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrThr 340
QY 1227 CCGATCCACTGGGAGCAGGAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGlnIleSerValGlyLeu 352

RESULT 4

AAW23835

ID AAW23835 standard; Protein: 352 AA.

XX

AC AAW23835;

XX

DT 08-JUN-1998 (first entry)
XX Human CC chemokine receptor 5 (CCR5).
XX CC chemokine receptor 5; CCR5; G-protein coupled receptor;
KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;
KW transgenic animal.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 29..55
FT Domain /label= I
FT /note= "transmembrane domain"
FT Region 109..120
FT Domain /note= "extracellular loop-1 (Claim 19)"
FT 104..126
FT /label= III
FT /note= "transmembrane domain"
FT 143..171
FT /label= IV
FT /note= "transmembrane domain"
FT Region 187..210
FT /note= "extracellular loop-2 (Claim 19)"
FT Domain 194..219
FT /label= V
FT /note= "transmembrane domain"
FT 238..258
FT /label= VI
FT /note= "transmembrane domain"
FT Region 261..276
FT /note= "extracellular loop-3 (Claim 19)"
FT Domain 277..300
FT /label= VII
FT /note= "transmembrane domain"
XX
XX WO9745543-A2.
XX
XX 04-DEC-1997.
XX
XX 28-MAY-1997; 97WO-US09586.
XX
XX 28-MAY-1996; 96US-0018508.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Alkhatib G, Berger EA, Broder CC, Combadiere C;
XX Feng Y, Kennedy PE, Murphy PM;
XX WPI: 1998-032650/03.
XX N-PSDB; AAT76920.
XX
XX CC chemokine receptor 5 polypeptide - used to inhibit membrane
XX fusion between HIV and a target cell
XX
XX Claim 68; Fig 1C; 70pp; English.
XX
XX This protein sequence comprises of a novel human macrophage-selective
XX CC chemokine receptor that has been designated CCR5. The sequence
XX was deduced from an isolated cDNA clone (see AAT76920). An Alal27Leu
XX variant (see W238340 of CCR5 was also identified. The susceptibility
XX of human macrophages to HIV infection depends on cell surface
XX expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane
XX superfamily of G-protein coupled cell surface molecules. It plays
XX an essential role in the membrane fusion step of infection by some
XX HIV isolates. The establishment of stable, non-human cell lines
XX and transgenic mammals having cells that coexpress human CD4 and
XX CCR5 provides valuable tools for research of HIV infection.
XX Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding
XX agents capable of blocking membrane fusion between HIV and target
XX cells represent potential anti-HIV therapeutics for macrophage
XX tropic strains of HIV.

SQ Sequence 352 AA;
Alignment Scores: 3.38e-187 Length: 352
Pred. No.: 1745.50 Matches: 340
Score: 96.60% Conservative: 1
Best Similarity: 96.32% Mismatches: 0
Best Local Similarity: 66.42% Indels: 12
Query Match: 19 Gaps: 1
DB:
US-09-938-703-3 (1-1442) x AAW23835 (1-352)
QY 240 ATGGATTATCAAGTGTCAAGTCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspIyrGlnValSerSerProIleIyrAspIleasnIyrIyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAGCAAAATCGCCGCGCTGCTGCTCCGCTCTACTCTACTGGTG 359
Db 21 GlnIysIleasnValIysGlnIleAlaAlaArgLeuLeuProLeuIyrSerLeuVal 40
QY 360 TTCATCTTGGTTCGCGCAACATGCTGCTCATCTCTCATCTCTGATAAATCGCAAGG 419
Db 41 PheIlePheGlyPheValIysMetLeuValIleIleIleIleIleIleIleIleIle 60
QY 420 CTGAAGAGCATGACTGACATCTACTGCTCAACTGCGCATCTCTGACCTGTTTTCTT 479
Db 61 LeuIysSerMetThrAspIleIyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTCCCTGGGCTCACTATGTCGCCGCGAGTGGGACTTTGGAAATACATG 539
Db 81 LeuThrValProPheTrpAlaHisIyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTGTGACAGGCTCTATTTATAGGCTCTTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrClyLeuIyrPheIleGlyPheSerGlyIlePheIleIle 120
QY 600 CTCCTGCAATCGATAGGTACTGCTGCTGCTCATGCTGTGTTGCTTAAAGCCAGG 659
Db 121 LeuLeuThrIleAspArgIyrLeuAlaValAlaValAlaValPheAlaLeuIlysAlaArg 140
QY 660 AGGTTCACCTTTGGGTGTGACAGTGTGATCATCTGGTGGTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCACAGGATCATCTTACCAGATCTCAAAAGAGGTCTTCAATACACCTGACAGTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnIysGluGlyLeuHisIyrThrCysSer 180
QY 780 CATTTTCATAC-----ATTAAGATAGTACAT 806
Db 181 HisPheProIyrSerGlnIyrGlnPheTrpIysAsnPheGlnThrLeu-IysIleValI 200
QY 807 CTTGGGGTGGTCTGCGCTGCTGTCATGTCATCTGCTACTCGGGAATCTTAAAC 866
Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysIyrSerGlyIleLeuIysTh 220
QY 867 TCTGCTTCGGTGTGCAATAGACAGACAGGCTGTGAGGCTGTATCTTCCACAT 926
Db 220 rLeuLeuArgCysArgAsnGluIysIysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGTTTATTTCTTCTTCTGGGTGCTCCCTACACATGTCCTCTCTGACACCTT 986
Db 240 eMetIleValIyrPheLeuPheTrpAlaProIyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTGGCTGAAATGAAATGAGTACAGTCTTAACAGTGTGGCAACAGTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GTGACAGAGACTCTTGGATGACGACTGTGTCATCAACCCATCATCTATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysIleAsnProIleIleIyrAlaPheVa 300
QY 1107 CGGGGAGAGTTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGACATTCGCAACGCTT 1166

Db 300 lGlyGluPheArgAsnTyrLeuValPheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGAAATGCTTCTTATTTCCAGCAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrTh 340
 QY 1227 CCGATCCATCTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGlnGlnGluSerValGlyLeu 352

RESULT 5

AAW88232
 ID AAW88232 standard; Protein; 352 AA.

AC AAW88232;

XX 15-MAR-1999 (first entry)

XX HIV-1 co-receptor CCR5.

XX HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;
 KW gene therapy; human.

XX Homo sapiens.

XX Location/Qualifiers

XX 32..56

XX /note= "transmembrane domain 1"

XX 67..87

XX /note= "transmembrane domain 2"

XX Misc-difference 101

XX /note= "corresponds to TGT (Cys) in wild-type CCR5,

XX TGA (Stop) in CCR5m303"

XX 103..124

XX /note= "transmembrane domain 3"

XX 142..167

XX /note= "transmembrane domain 4"

XX 200..223

XX /note= "transmembrane domain 5"

XX 236..260

XX /note= "transmembrane domain 6"

XX 275..301

XX /note= "transmembrane domain 7"

XX W09854317-Al.

XX 03-DEC-1998.

XX 29-MAY-1998; 98WO-EP03437.

XX 30-MAY-1997; 97US-0048057.

XX (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.

XX Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;

XX WPI; 1999-059835/05.

XX DR N-PSDB; AAV84126.

XX New CCR5 variant protein of the HIV-1 co-receptor - useful in

XX developing resistance of CCR5-expressing cells to HIV-1 infection

XX Disclosure; Page 34-35; 55pp; English.

XX This is the amino acid sequence of wild-type human CCR5, which
 CC serves as a co-receptor for infection by macrophage-tropic
 CC (M-tropic) strains of HIV-1. The invention relates to the
 CC identification of a CCR5 variant (see AAW88232), designated CCR5m303,
 CC comprising the first two transmembrane domains of wild-type CCR5,
 CC but lacking transmembrane domains 3-7. The presence of the
 CC CCR5m303 variant with the wild type CCR5 allele shows a positive
 CC correlation with resistance to infection with M-tropic HIV-1

CC strains, and may indicate slower progression of the disease. The
 CC detection of CCR5 variants may be used to identify individuals at
 CC lower risk of infection relative to the general population who, if
 CC infected, may exhibit slower progression to AIDS. Probes and
 CC primers (see AAW84127-36) are provided for use in diagnostic methods
 CC for detecting the presence of such variants. A method is provided
 CC for inhibiting HIV-1 infection of a cell expressing the CCR5
 CC receptor. This involves introducing a nucleic acid encoding a CCR5
 CC variant into the cell, thereby reducing the number of functional
 CC CCR5 molecules present on the cell surface.

XX Sequence 352 AA;

Alignment Scores:

Pred. No.: 3.38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 20 Gaps: 1

US-09-938-703-3 (1-1442) x AAW88232 (1-352)

QY 240 ATGGATTATCAAGTCTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleYrAspIleAsnTyrThrSerGluProCys 20
 QY 300 CAAAAATCAATGTGAAGCAATCGCAGCCGCTCCTCGCTCGCTACTACTACTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAargLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGTTTGTGGCAACATGCTGGTCACTCCTCATCCTGATAAATGCAAAAGG 419
 Db 41 PheIlePheGlyValGlyAsnMetLeuValIleLeuIleLeuLeuAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCATCTCTGACCTGTTTCTCT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCGCTGTTGTTTAAAGCCAGG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGACAAATCGATAGTACCTGGCTCTGCTCCATGCTGTGTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTTGGGTTGGTGACAGTGTGATCAGTGGGTTGGGTTGGTGGTGGTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCACAGGAATCATCTTACAGATCTCAAAAGAGGCTTTCATACACCTCAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTCCTCATAC-----ATTAAGATAGTATCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeuLysIleValI 200
 QY 807 CTTGGGCTGTGCTCTCCGCTGCTGCTCATGCTGCTCATGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCGGTCTCGAAATGAGAGAGAGCAGCAGGCTGTGAGGCTTATCTTCCATCAT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTGTTTATTTCTCTCTGGGCTCCCTACACATGTCCTCTCTCTCTCTCTCTCT 986

Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CGAGAAATCTTTGGCTGAATATTCGAGTACCTCTACAGAGTTGGACCAAGTATGCA 1046
 Db 260 eGlnGluPheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG1 280
 QY 1047 GTGCAGAGACTCTTGGATGACGACTGCTGCATCAACCCATCATCTATGCTTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGTCTGAGAACTACCTCTTACTTCTTCCAAAGACATGTCGCAACGCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGCTCTATTTCACAGAGAGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATCCACTGGGAGCAGGAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 6
 AAG80111
 ID AAG80111 standard; Protein: 352 AA.
 AC AAG80111;
 XX
 DT 17-JAN-2002 (first entry)
 DE Human CCR5 protein.
 KW Chemokine; tumour diagnosis; colorectal; prostatic; organ rejection;
 KW inflammation; autoimmune disease; metastasis; bronchial asthma; lupus;
 KW chronic bowel inflammation; rheumatoid arthritis; cytostatic;
 KW antiinflammatory; antischismatic; immunosuppressive; dermatological;
 KW antirheumatic; antiarthritic.
 OS Homo sapiens.
 XX
 PN W0200172830-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 02-APR-2001; 2001WO-EP03708.
 XX
 PR 31-MAR-2000; 2000DE-1016013.
 XX
 PA (IPFP-) IPF PHARM GMBH.
 PA (FORS/) FORSMANN U.
 XX
 PI Forssmann W, Adermann K, Heitland A, Spodsborg N;
 XX
 DR WPI; 2001-626256/72.
 XX
 PT Diagnostic agent containing two or more receptor-specific ligands,
 PT useful for detecting tumors, inflammation etc., also therapeutic use of
 PT ligand inhibitors -
 XX
 PS Disclosure; Page 10; 26pp; German.
 XX

CC This invention describes a novel diagnostic agent (A) comprising at least
 CC two different ligands (I) for receptors (II) that are implicated in
 CC disease. (A) are used for the diagnosis of tumors (especially colorectal
 CC or prostatic), organ rejection, inflammation and autoimmune diseases.
 CC Also inhibitors of (I) are used therapeutically against tumors (and their
 CC metastases), inflammation (particularly bronchial asthma or chronic bowel
 CC inflammation), or autoimmune diseases (rheumatoid arthritis or lupus),
 CC where the (cardio)vascular, lymphatic, respiratory, nervous, digestive,
 CC endocrine, motor or urogenital systems or skin are affected, and bone
 CC marrow diseases. The products of the invention are chemokine derivatives
 CC which have cytostatic, antiinflammatory, antiasthmatic,
 CC immunosuppressive, dermatological, antirheumatic, antiarthritic.

CC Chemokines act on specific tumor and inflammatory cells through a
 CC constellation of chemokine receptors (CR), which control migration and
 CC proliferation of these cells. AAG80045-AAG80128 represent human chemokine
 CC fragments used to illustrate the method of the invention.
 XX
 SQ Sequence 352 AA:
 Alignment Scores:
 Pred. No.: 338e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 22 Gaps: 1

US-09-938-703-3 (1-1442) x AAG80111 (1-352)

QY 240 ATGATTATCAAGTGTCAAGTCCAAATCATGACATCAATATTATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
 QY 300 CAAAAAATCAATGTGAAGCAAAATCCACCGCGCTCTCTCGCTCGCTCTACTCTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGTTTGTGGCAACATGCTGCTCATCTCTCATCTGATAAACTGCAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetIleValIleLeuIleAsnCysLysArg 60
 QY 420 CTGAGAGCATGACTGACTATCTACTGCTCAACCTGCCATCTCTGACCTGTTTTCTCT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCAGTGGACTTGGAAATACAAG 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCTGACAAATCGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCACTTTGGGTTGGTGCACAGTGTGATCTTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTTACCAGATCTCAAAAAGAGGTCTTCTTATACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
 QY 807 CTTGGGGCTGTCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
 QY 867 TCTGCTCGGTGCTGCAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTACCATT 926
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTTTATTTTCTCTCTGCGGCTCCCTACAACTGCTCTCTCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATTTCTTTGGCTGTAATAATTGACAGTAGCTCTAACAGGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG1 280

Amino acids 2-18 in the N-terminal region of CCR5 comprise an HIV-1 gp120-binding site that determines the specificity of the interaction between CCR5 and HIV-1 gp120. Post-translational sulfation of the tyrosine residues in the CCR5 N-terminus is required for gp120 binding and may critically modulate the susceptibility of target cells to HIV-1 infection *in vivo*. The invention provides sulfated peptides (see AAB82947) that are based on the CCR5 N-terminal region and which are effective for inhibiting HIV-1 binding to CCR5. These peptides are used in claimed methods of inhibiting HIV infection of CD4⁺ cells, of preventing CD4⁺ cells from becoming infected with HIV, of treating a subject whose CD4⁺ cells are infected with HIV, and of


```

Percent Similarity: 96.60%      Conservative: 1
Best Local Similarity: 96.32%    Mismatches: 0
Query Match: 66.42%             Indels: 12
DB: 22                          Gaps: 1

US-09-938-703-3 (1-1442) x AAF04321 (1-352)

```

QY 240 ATGGATTATCAAGTGTCAAGTCCAATCTATGACATCAATTATATACATCGGAGCCCTGC 299
 1 MetaspTyrGlnValSerSerProIleTyrAspIleasnTyrTyrThrSerGluProCys 20
 Db

QV 300 CAAAAATCAATGTGAAGCAAAATCGAGCCCGCTCCTGCTCCTACTCACTGGTG 359

Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProProLeuTyrSerLeuVal 40

360 TTCAATCTTTGGTTTGTGGGCAACATGCTGGTCATCCTCATCCTGATAAACTGCAAAAGG 419

Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60

QV 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479

Db 61 LeuLysSerMetThrAspIleTyrLeuIleuAsnIeuAlaIleSerAspIeuPheIeu 80

Qy 480 CTTACTGCCCTTCGGGCTCATTATGTCGCCGCCAGTGGGACTTTGGAAATACAAG 539

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2

[illegible]

PT TOT CYZATUPEENUNWOTUTYFENUTYFMEETREDTYFNEEWEDEBTUTYTTEEEHETTERETETE 120

F

2016年12月31日

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

2

[illegible]

18

| | | | |
|----|------|---|------|
| QY | 927 | CATGATTGGTTATTTCTCTCTCTGGGTCCTACACATATGTCTTCTCTCTGACACCTT | 986 |
| Db | 240 | eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh | 260 |
| QY | 987 | CCAGGAATTCCTTGGCCCTGAATAATTCAGTCAGTCTACAGGTGGACCAAGTATGCA | 104 |
| Db | 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetGI | 280 |
| QY | 1047 | GGTGACAGAGACTCTGGGATGACGCATGTCATCAACCCCATCATCTATGCGCTTGT | 1108 |
| Db | 280 | nValThrGluThrLeuGlyMetThrHisCysSylleAsnProIleLysrAlaPheVa | 300 |

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

QY 1047 GGTGACAGAGACTCTTGGATGACGACTGCTGCATCAACCCCATCATCTATGCGCTTGT 1106
 XX
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300

QY 1107 CGGGAGAGATTCAGAACTACCTCTAGTCTTCTCCAAAGCAGATGCCAAAGCCTT 1166
 XX
 Db 300 LGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320

QY 1167 CTGCAATGCTCTCTATTTTCCAGCAAGAGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 XX
 Db 320 ecysLysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340

QY 1227 CCGATCCTCTGGGAGCAGGAAATATCTCTGGGGTTG 1263
 XX
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 11
 AAM52828
 ID AAM52828 standard; Protein; 352 AA.
 XX
 AC AAM52828;
 XX
 DT 22-FEB-2002 (first entry)
 XX
 DE Human CC chemokine receptor 5 (CCR5).
 XX
 KW CCR5; CC chemokine receptor 5; human; HIV infection;
 KW human immunodeficiency virus; AIDS; acquired immunodeficiency syndrome;
 KW drug screening; identification.
 XX
 OS Homo sapiens.
 XX
 PN W0200171346-A2.
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09155.
 XX
 PR 21-MAR-2000; 2000US-190946P.
 PR 21-MAR-2000; 2000US-190996P.
 PR 21-MAR-2000; 2000US-191299P.
 PR 20-MAR-2001; 2001US-0813448.
 PR 20-MAR-2001; 2001US-0813651.
 PR 20-MAR-2001; 2001US-0813653.
 XX
 PA (CONS-) CONSENSUS PHARM INC.
 XX
 PI Nestor JJ, Wilson CJ, See RH, Tan Hehir CA;
 XX
 DR WPI: 2002-010610/01.
 DR N-PSDB; ARA02317.
 XX
 XX Identifying CC chemokine receptor 5 binding compound for treating AIDS,
 PT comprises binding a molecule from library to a molecule having binding
 PT property corresponding to CCR5 and identifying bound molecule -
 XX
 XX Example 3; Fig 4a; 50pp; English.
 PS
 XX The invention relates to a method for identifying a binding compound
 CC for CC chemokine receptor 5 (CCR5). The method involves screening a
 CC library of test molecules (particularly peptides) with immobilised CCR5,
 CC and then identifying those molecules which bind. The invention also
 CC relates to CCR5-binding molecules identified using the method of the
 CC invention, methods for identifying consensus motifs for CCR5-binding
 CC peptides, a transfer vector encoding tagged CCR5, a computer-aided
 CC methods for determining the relative binding affinity of a test molecule
 CC to CCR5 and a computer aided drug screening assay that utilises the
 CC three-dimensional structure of CCR5. Compounds identified using the
 CC methods of the invention are useful for treating or preventing HIV
 CC (human immunodeficiency virus) infection or AIDS (acquired
 CC immunodeficiency syndrome) in a patient. The methods of the invention
 CC may also be used to identify agonists or antagonists of the interaction
 CC of CCR5 with its natural ligand, and to determine a binding motif for

CC CCR5. The present sequence represents human CCR5.

XX
 SQ Sequence 352 AA;

Alignment Scores:

Pred. No.: 3,38e-187 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 Gaps: 23

US-09-938-703-3 (1-1442) x AAM52828 (1-352)

QY 240 ATGGATTATCAAGTGTCAGAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20

QY 300 CAAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCGCTCTACTACTACTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40

QY 360 TTCATCTTTTGTGGTGGCAACATGCTGCTCATCTCTCATCTCTGATAAATGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60

QY 420 CTGAGAGACATGACTGACATCTACTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80

QY 480 CTTACTCTCCCTCTCTGGGCTCCTACTATCTGCGCCGCGAGTGGGACTTTGGAATACAATG 539
 Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaGlnTyrPaspPheGlyAsnThrMet 100

QY 540 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTCTCGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleLe 120

QY 600 CTCCTGACAAATCGATAGTACCTGGCTCTGCTCATGCTGCTGCTTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140

QY 660 ACGGTCACCTTTGGGTGGTGCAAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTyrValValAlaValPheAlaSer 160

QY 720 CTCACAGGAATCATCTTACAGACTCTCAAAAGAGGCTTCTCATACCTGACCTGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180

QY 780 CATTTTCCATAC-----ATTAAGAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValI 200

QY 807 CTTGGGGCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220

QY 867 TCTGCTTCGCTGCTGCAATGAGAAGAGGACAGGCTGAGGCTTATCTTCACCAT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240

QY 927 CATGATTGTTTATTTTCTCTCTCTGGGCTCCCTACAACATGCTCTCTCTCTCTCTCT 986
 Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuAsnThrPh 260

QY 987 CCAGGAATCTTTGGGCTGAAATATTCAGTAGTCTTAACAGGTTGGACAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280

QY 1047 GGTGACAGAGACTCTTGGATGACGACTGCTGCATCAACCCCATCATCTATGCGCTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300

QY 1107 CGGGAGAGTTCAGAACTACCTTTAGTCTTCTTCCAAAGACATGCGCAAGCGCTT 1166
Db 300 IGIyGLuLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTTCTTATTTCCAGCAAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGGAATAATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGlnIleSerValGlyLeu 352
RESULT 12
ID AAY41280
ID AAY41280 standard; Protein; 439 AA.
XX AAY41280;
AC AAY41280;
XX 11-FEB-2000 (first entry)
XX Fusion protein containing human chemokine receptor CCR-5.
DE Prostaglandin; PG; E2EP3 receptor; E2EP2 receptor; CCR-5; human;
KW chemokine receptor; ss2 adrenergic receptor; small G-protein rho;
KW renal outer medullary potassium ion channel protein; ion-channel protein;
KW lambda phage repressor protein; G-protein coupled receptor; bacteria;
KW biochemical; vaccine; immunohistochemical; orphan receptor; HIV.
XX Synthetic.
OS Homo sapiens.
XX W0953033-AL.
PN 21-OCT-1999.
XX 16-APR-1999; 99WO-US08214.
XX 16-APR-1999; 98US-0081989.
XX (UYVA-) UNIV VANDERBILT.
XX Breyer RM, Ma L, Kennedy C;
XX WPI: 1999-620416/53.
XX N-PSDB; AA224738.
XX New nucleic acid constructs for high level expression of eukaryotic
PT proteins in bacteria, for producing e.g. chemokine receptor CCR-5 for
PT preventing HIV infection
XX Disclosure; Page 49-53; 81pp; English.
XX The invention provides isolated nucleic acid sequences that encode rabbit
CC prostaglandin (PG) E2EP3 receptor, human PG E2EP2 receptor, human
CC chemokine receptor CCR-5, human ss2 adrenergic receptor, rat renal outer
CC medullary potassium ion channel protein or human small G-protein rho,
CC together with deduced protein sequences. Also provided is a method for
CC the production of eukaryotic proteins by culturing bacteria transformed
CC with vectors containing the above nucleic acid sequences or a nucleic
CC acid (I) that comprises: (i) first sequence that encodes either a
CC sequence comprising at least three positively charged amino acids, or a
CC DNA-binding protein, or a lambda phage repressor protein, placed upstream
CC of, and in frame with, (ii) a sequence encoding a protein. (I) are used
CC for recombinant production of eukaryotic proteins, particularly membrane
CC proteins, G-protein coupled receptors or ion-channel proteins, in
CC bacteria. These proteins are useful for biochemical or structural
CC studies; as therapeutic agents; in diagnostic and screening assays and as
CC antigens for use in vaccines, and for raising antibodies that are useful
CC as immunohistochemical markers, e.g. for orphan receptors or ion
CC channels. Antibodies raised against the chemokine receptor CCR-5 can be
CC used (when administered as antiserum or generated in vivo) to prevent
CC entry of human immune deficiency virus (HIV) into cells.

XX Sequence 439 AA;
SQ Alignment Scores: 3.72e-187 Length: 439
Pred. No.: 1745.50 Matches: 340
Score: 96.60% Conservative: 0
Percent Similarity: 96.32% Mismatches: 0
Best Local Similarity: 66.42% Indels: 12
Query Match: 20 Gaps: 1
DB: 1
US-09-938-703-3 (1-1442) x AAY41280 (1-439)
QY 240 ATGGATTATCAAGTGTCAGTCCATCTATGACATCAATTTATATACATGGAGCCCTGC 299
Db 79 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 98
QY 300 CAAAAATCAATGTGAAGCAAAATCGACCCCGCTCCTCGCTCGCTACTACTACTGCTG 359
Db 99 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProProLeuTyrSerLeuVal 118
QY 360 TTCATCTTTGGTTTGTGGGCAACATGCTGTCATCTCTCATCTCTGCTGCTGCTGCTG 419
Db 119 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 138
QY 420 CTGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 139 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 158
QY 480 CTTACTGTCCTTCGCTGGGCTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
Db 159 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 178
QY 540 TGTCACTCTTGACAGGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
Db 179 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 198
QY 600 CTCTCTGACAAATCGATAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 199 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 218
QY 660 ACGTGCACCTTTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
Db 219 ThrValThrPheGlyValValThrSerValIleThrTyrValIleValAlaValPheAlaSer 238
QY 720 CTCCAGGAATCATCTTTACAGATCTCAAAAAGAGGCTCTCATACACCTGCAGCTCT 779
Db 239 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 258
QY 780 CATTTCCTCATC-----ATTAAGATAGTCAT 806
Db 259 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 278
QY 807 CTTGGGGCTGGTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
Db 278 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 298
QY 867 TCTGCTTCGGTGTCAAAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTACCAT 926
Db 298 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrI 318
QY 927 CATGATGTTTATTTCTCTCTGGGCTCCCTACACATTTGCTCTCTCTCTCTCTCTCTCT 986
Db 318 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 338
QY 987 CCAGGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
Db 338 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 358
QY 1047 GGTGACAGAGACTCTTGGGATGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
Db 358 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 378

AC AAM52829;
XX 22-FEB-2002 (first entry)
DT Human CCR5 Gln 55 variant.
XX
XX CCR5; CC chemokine receptor 5; human; HIV infection;
KW human immunodeficiency virus; AIDS; acquired immunodeficiency syndrome;
KW drug screening; identification; variant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 55 /note= "Glu replaces wild-type Leu; encoded by CNG"
FT Misc-difference 58 /note= "Encoded by AGC"
FT
XX W0200171346-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09155.
XX
XX 21-MAR-2000; 2000US-190946P.
XX 21-MAR-2000; 2000US-190996P.
XX 21-MAR-2000; 2000US-191299P.
XX 20-MAR-2001; 2001US-081344B.
XX 20-MAR-2001; 2001US-0813651.
XX 20-MAR-2001; 2001US-0813653.
XX
XX (CONS-) CONSENSUS PHARM INC.
XX
XX Nestor JJ, Wilson CJ, See RH, Tan Behir CA;
PI
XX
XX WPI: 2002-010610/01.
XX N-PSDB; ABR02318.
XX
XX Identifying CC chemokine receptor 5 binding compound for treating AIDS,
XX comprises binding a molecule from library to a molecule having binding
XX property corresponding to CCR5 and identifying bound molecule -
XX
XX Example 3; Fig 4B; 50pp; English.
XX
XX The invention relates to a method for identifying a binding compound
XX for CC chemokine receptor 5 (CCR5). The method involves screening a
XX library of test molecules (particularly peptides) with immobilised CCR5,
XX and then identifying those molecules which bind. The invention also
XX relates to CCR5-binding molecules identified using the method of the
XX invention, methods for identifying consensus motifs for CCR5-binding
XX peptides, a transfer vector encoding tagged CCR5, a computer-aided
XX methods for determining the relative binding affinity of a test molecule
XX to CCR5 and a computer aided drug screening assay that utilises the
XX three-dimensional structure of CCR5. Compounds identified using the
XX methods of the invention are useful for treating or preventing HIV
XX (human immunodeficiency virus) infection or AIDS (acquired
XX immunodeficiency syndrome) in a patient. The methods of the invention
XX may also be used to identify agonists or antagonists of the interaction
XX of CCR5 with its natural ligand, and to determine a binding motif for
XX CCR5. The present sequence represents a naturally occurring variant of
XX human CCR5 in which there is a glutamine, rather than a leucine, at
XX position 55.
XX
XX SQ Sequence 352 AA;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 1.6e-186 | Length: | 352 |
| Score: | 1739.50 | Matches: | 339 |
| Percent Similarity: | 96.32% | Conservative: | 1 |
| Best Local Similarity: | 96.03% | Mismatches: | 1 |
| Query Match: | 66.19% | Indels: | 12 |
| DB: | 23 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x AAM52829 (1-352)

QY 240 ATGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATACATCGAGCCCTGC 299
DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTAAGCAAAATCGAGCCCTCTCTGCTCGCTCTACTACTCTGGTG 359
DB 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGTGGCAACATGCTGGTCACTCTCATCTCTGATAAAGCAAGG 419
DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleGlnIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACTGACATCTACCTCTCAACCTGGCCATCTCTGACCTGTCTTCTT 479
DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGGCGCCAGTGGGACTTTGGAAATACAATG 539
DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTTCTCTGGAATCTTCTCATATC 599
DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGCAATCGATAGGTACCTGGCTGCTCATCTGTGTGTCTTTAAAGCCAGG 659
DB 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 AGGTCACCTTTGGGTTGGTCAAGTGTGATCCTCTGGTGGTGGTGGTGGTGGTCT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCOCAGGAATCATCTTTACAGATCTCAAAAAGAGGCTCTTACATCTGCACCTCT 779
DB 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCCTCATAC-----ATTAAAGATAGTCAT 806
DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
QY 807 CTTGGGGCTGGTCTCGCTGCTTGTCTCATGCTCATCTCTACTCTGGAATCTCTAAAAAC 866
DB 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTGGTGTCTGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
DB 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATTGTTATTTCTCTCTCTGGGCTCCCTCAACATCTCTCTCTCTCTCTCTCTCT 986
DB 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATTTCTTGGCTGAAATTAATTCAGTAGCTCTACAGGTGGACCAAGCTATGCA 1046
DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTTCAGAGAGATCTCTGGATGAGCAGTCTGTCATCAACCCCATCATCTATGCTTTGT 1106
DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAAAGTTTCAGAACTACCTCTTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1166
DB 300 lGlyGluLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAAAATGCTTCTTATTTTCCAGAGAGGCTCCCGAGCGAGCAAGCTCAGTTACAC 1226
DB 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGGAGGAAATATCTGTGGCTTG 1263
DB 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 15

AAE07039

ID AAE07039 standard; Protein; 352 AA.

XX AC AAE07039;

XX DT 16-OCT-2001 (first entry)

XX DE Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.

XX KW Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;

XX KW human immunodeficiency virus; antimicrobial; vasodilator; vulnerary;

XX KW cytotactic; immunosuppressive; nontropic; neuroprotective; gene therapy;

XX KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;

XX KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;

XX KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;

XX KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;

XX KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;

XX KW cardiovascular disorder; myocardial ischaemia.

XX OS Homo sapiens.

XX PW WO200158915-A2.

XX PD 16-AUG-2001.

XX PF 09-FEB-2001; 2001WO-US04152.

XX PR 09-FEB-2000; 2000US-0181258.

XX PR 09-MAR-2000; 2000US-0187999.

XX PR 22-SEP-2000; 2000US-0234336.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Roschke V, Li Y, Ruben SM;

XX DR WPI; 2001-489865/53.

XX DR N-PSDB; AAD13198.

XX PT Isolated nucleic acid encoding a human G-protein chemokine receptor

XX PT (CCR5) HDGNR10 polypeptide, useful for preventing or treating

XX PT autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative

XX PT disorders and neurodegenerative disorders -

XX PS Example 40; Page 486-487; 495pp; English.

XX CC The present sequence is human G-protein chemokine receptor (CCR5) HDGNR10

XX CC protein. CCR5 HDGNR10 antibodies are useful for treating, preventing

XX CC or ameliorating a disease or disorder associated with inflammation,

XX CC defective or aberrant chemotaxis of immune cells, HIV infection (such as

XX CC Pneumocystis carinii pneumonia or Kaposi's sarcoma) or defective or

XX CC aberrant T-cell antigen presenting cell interaction. The disease or

XX CC disorder may also be an infectious disease (e.g. a viral infection such

XX CC as an early stage HIV infection, a cytomegalovirus infection, or a

XX CC poxvirus infection) an autoimmune disease (e.g. rheumatoid arthritis) or

XX CC a neurodegenerative disorder. The disease or disorder may be associated

XX CC with aberrant CCR5 expression, lack of CCR5 function, aberrant CCR5

XX CC ligand expression, or lack of CCR5 ligand function. CCR5 HDGNR10 protein

XX CC is used as a food additive or preservative to increase or decrease

XX CC storage capabilities. CCR5 HDGNR10 DNA are useful for chromosome

XX CC identification and in gene therapy. CCR5 HDGNR10 DNA, protein,

XX CC antibodies, agonists and antagonists are also useful in the

XX CC diagnosis, treatment and prevention of cancer (breast, ovary, adrenal

XX CC gland, bone, bone marrow, gastrointestinal tract, liver, lung,

XX CC urogenital); immune disorders (Addison's disease, allergies, autoimmune

XX CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

XX CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative

XX CC colitis); cardiovascular disorders (myocardial ischaemia) and

XX CC wound healing.

XX CC Sequence 352 AA;

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

| | | | |
|---|-----------|---|------|
| Alignment Scores: | 2.69e-186 | Length: | 352 |
| Pred. No.: | 1737.50 | Matches: | 339 |
| Score: | 96.32% | Conservative: | 1 |
| Percent Similarity: | 96.03% | Mismatches: | 1 |
| Best Local Similarity: | 66.11% | Indels: | 12 |
| Query Match: | 22 | Gaps: | 1 |
| DB: | | | |
| US-09-938-703-3 (1-1442) x AAE07039 (1-352) | | | |
| QY | 240 | ATGATTATCAAGTGTCAAGTCCCAATCATATGACATCAATATATATACATGAGGCCCTGC | 299 |
| DB | 1 | MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys | 20 |
| QY | 300 | CAAAATAATCAATGTGAAGCAATCCAGCCCGCTCCTCGCTCGCTCTACTCATGTGGT | 359 |
| DB | 21 | GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal | 40 |
| QY | 360 | TTTCATCTTTGGTTTGGTGGCAACATGCTGTCATCTCATCTCATCTGATAAACTCAAAAGG | 419 |
| DB | 41 | PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuAsnCysLysArg | 60 |
| QY | 420 | CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT | 479 |
| DB | 61 | LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu | 80 |
| QY | 480 | CTTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCAGTGGACTTGGAAATACATG | 539 |
| DB | 81 | LeuThrValProPheThrAlaHisTyrAlaAlaIleThrPaspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCAACTCTTGACAGGGCTCTATTTATAGGCTTCTCTCTGGAATCTCTTTCATCATC | 599 |
| DB | 101 | CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyPhePheIleIle | 120 |
| QY | 600 | CTCCTGACAATGATAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 659 |
| DB | 121 | LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg | 140 |
| QY | 660 | ACGTCACCTTTGGGTGGTGACAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGT | 719 |
| DB | 141 | ThrValThrPheGlyValValThrSerValIleThrThrValValAlaValPheAlaSer | 160 |
| QY | 720 | CTCCAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCATTCACCTGACGCTCT | 779 |
| DB | 161 | LeuProGlyIleIlePheThrArgSerGlnLysGlyLeuHisTyrThrCysSerSer | 180 |
| QY | 780 | CAITTTCCATAC-----ATTAAAGATAGTCAT | 806 |
| DB | 181 | HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValI | 200 |
| QY | 807 | CTTGGGCTGGTCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 866 |
| DB | 200 | eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh | 220 |
| QY | 867 | TCTGCTTCGGTGTGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG | 926 |
| DB | 220 | rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI | 240 |
| QY | 927 | CATGATTTGTTTATTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 986 |
| DB | 240 | eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh | 260 |
| QY | 987 | CCAGGAATCTTTGGCTGAAATATTCAGTAGCTCTAACAGGTGGACCAAGCTATGCA | 1046 |
| DB | 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGl | 280 |
| QY | 1047 | GTCACAGAGACTCTTGGGATGAGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCG | 1106 |
| DB | 280 | nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa | 300 |
| QY | 1107 | CGGGGAGAGCTTCAGAACTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT | 1166 |
| DB | 300 | IcIyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh | 320 |

| | | | |
|----|------|--|------|
| QY | 1167 | CTGCAAAATGCTGTTCTATTTCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACAC | 1226 |
| Db | 320 | eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh | 340 |
| QY | 1227 | CCGATCCACTGGGGAGCAGGAATATCTGTGGGCTTG | 1263 |
| Db | 340 | rArgSerThrGlnGlnGlnIleSerValGlyLeu | 352 |

Search completed: June 3, 2003, 18:45:05
Job time : 91.6348 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_p2p model.

Run on: June 3, 2003, 18:41:24 ; Search time 6.73177 Seconds
(without alignments)
5754.334 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequence: 1 ATTAAGATAGTATCTTGG.....CTCTGCTGGTGGTCAAAAT 94

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09938703/runat_03062003_161407_22827/app_query.fasta.1.1870
-DB=SPREMBL_21 -OPMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.ecgn.1.1.192.etrnat_03062003_161407_22827 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match Length | ID | Description |
|------------|-------------|--------------|-------|--------------------|
| 1 | 164 | 100.0 | 215 4 | 075303 homo sapien |

| ID | 075303 | PRELIMINARY; | PRT; | 215 AA. |
|----|--|--------------|------|---------|
| AC | 075303; | | | |
| DT | 01-NOV-1998 (TRENBLrel. 08, Created) | | | |
| DT | 01-NOV-1998 (TRENBLrel. 08, Last sequence update) | | | |
| DE | 01-DEC-2001 (TRENBLrel. 19, Last annotation update) | | | |
| DE | CC-chemokine receptor. | | | |
| GN | CCR-5. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Tse L., Ehrenberg P.K., Chang G., Michael N.L.; | | | |
| RT | Genomic Organization and Functional Characterization of the Complete | | | |
| RT | Transcription Unit for the Chemokine Receptor CCR-5, a Major Entry Co- | | | |
| RL | Receptor for HIV-1.; | | | |
| RL | Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AF009962; AAC23944.1; - | | | |
| DR | InterPro; IPR000276; GPCR_Rhodpsn. | | | |
| DR | Pfam; PF00001; 7tm1; 1. | | | |
| DR | PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. | | | |
| DR | PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1. | | | |

ALIGNMENTS

RESULT 1
075303
ID 075303 PRELIMINARY; PRT; 215 AA.
AC 075303;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CC-chemokine receptor.
GN CCR-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tse L., Ehrenberg P.K., Chang G., Michael N.L.;

KW Receptor.
SQ SEQUENCE 215 AA; 23946 MW; 3C9146C76BA416F7 CRC64;

Alignment Scores:

Pred. No.: 4,51e-16 Length: 215
Score: 164.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q75303 (1-215)

QY 1 ATTAAGATAGTCATCTGGCGTCCCTGGCTGCTGTCATGTCATCTGCTACTC 60
Db 185 ILLEYSASPSEHISLEUGLYALAGLYPROALLAALACYSHISGLYHISLEULEU 204
QY 61 GGGAATCTCTAAACTCTGCTCGTGCAGAA 93
Db 205 GLYASNPOLYSANSEALASERVALSERIYS 215

RESULT 2

Q9P1T4
ID Q9P1T4 PRELIMINARY; PRT; 60 AA.

AC Q9P1T4
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Chemokine receptor CCR5 (Fragment).
GN CCR5.

OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]

RN RN
RP SEQUENCE FROM N.A.

RX MEDLINE-99392755; PubMed=10465086;
RA Aikhionbare F.O., Newman C., Womack C., Roth W.W., Stringer H.G. Jr.,
RA Bond V.C.;
RT "Characterization of a third CCR5 amplicon from CCR5-delta32-
RT heterozygous HIV-1-infected individuals.";
RL AIDS 13:1585-1586(1999).

DR EMBL; AF056020; AAF65579.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_FL_2; 1.

KW Receptor.
FT NON_TER 1 1
FT NON_TER 60 60
SQ SEQUENCE 60 AA; 7156 MW; AFF4B9CAF6B80AFB CRC64;

Alignment Scores:

Pred. No.: 2,81e-14 Length: 60
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9P1T4 (1-60)

QY 2 TTAAGATAGTCATCTGGCGTCCCTGGCTGCTGTCATGTCATCTGCTACTCG 61
Db 22 LEULYSLEVALLEUGLYLEUVALLEUVALLEUVALLEUVALLEUVALLEUVAL 41

QY 62 GGAATCTCTAAACTCTGCTCGTGCAGAAAT 94

Db 42 GLYLEULEULYSLEUVALLEUVALLEUVALLEUVALLEUVALLEUVALLEU 52

RESULT 3

O14694
ID O14694 PRELIMINARY; PRT; 333 AA.

AC O14694;

DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CCR5 receptor (Fragment).
GN CCR5.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]

RN RN
RP SEQUENCE FROM N.A.

RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;

RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).

DR EMBL; AF011504; AAB65704.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.

DR PROSITE; PS00262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER 333 333

SQ SEQUENCE 333 AA; 38174 MW; AEFBA07A67893AEB CRC64;

Alignment Scores:

Pred. No.: 3e-14 Length: 333
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x O14694 (1-333)

QY 2 TTAAGATAGTCATCTGGCGTCCCTGGCTGCTGTCATGTCATCTGCTACTCG 61
[1]

Db 177 LEULYSLEVALLEUGLYLEUVALLEUVALLEUVALLEUVALLEUVALLEUVAL 196

QY 62 GGAATCTCTAAACTCTGCTCGTGCAGAAAT 94

Db 197 GLYLEULEULYSLEUVALLEUVALLEUVALLEUVALLEUVALLEUVALLEU 207

RESULT 4

Q9UN28
ID Q9UN28 PRELIMINARY; PRT; 339 AA.

AC Q9UN28;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]

RN RN
RP SEQUENCE FROM N.A.

RA Kunstman K., Chen Z., Korber B., Oprendeck J., Stanton J., Agly M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.,
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF161913; AAD47670.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.

DR PROSITE; PS00262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER 1 1

FT NON_TER 339 339

SQ SEQUENCE 339 AA; 39183 MW; 56C819F92D6DB1A6 CRC64;

| | |
|--|---|
| Alignment Scores: | |
| Pred. No.: | 3e-14 |
| Score: | 152.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 92.68% |
| DB: | 4 |
| US-09-938-703-3_COPY_792_885 (1-94) x Q9UN28 (1-339) | |
| QY | 2 TTAAGATGCTCATCTGGCGTGTCTGCCGCTGCTGTGTCATGG |
| Db | 189 LeuNlysleValIleLeuGlyLeuValLeuProLeuLeuValMetv |
| QY | 62 GGAATCCTAAAAACTCTGCTCGGTGTCGAAAT 94 |
| Db | 209 GlyIleLeuLysThrLeuLeuArgCysArgasn 219 |
| RESULT 5 | |
| Q9UN27 | PRELIMINARY; PRT; 339 AA. |
| ID | Q9UN27 |
| AC | Q9UN27 |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last sequence update) |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) |
| DE | C-C chemokine receptor 5 (Fragment). |
| GN | Homo sapiens (Human). |
| OS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute- |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; HC |
| OX | NCBI_TaxID=9606; |
| [1] | SEQUENCE FROM N.A. |
| RN | Kunstanm K., Chen Z., Korber B., Oprondek J., Stanton J. |
| RA | Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., W |
| RA | "Sequences of the CCR5 genes from diverse simian and pr |
| RT | species."; |
| RL | Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases |
| DR | EMBL; AF161914; RAD47671.1; |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. |
| DR | Pfam: PF00001; 7tm_1.1. |
| DR | PRINTS; PR00237; GPCRRHODPSN. |
| DR | PROSITE; PS00347; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. |
| DR | PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1. |
| KW | Receptor. |
| FT | NON_TER 1 1 |
| FT | NON_TER 339 339 |
| SQ | SEQUENCE 339 AA; 39086 MW; 88AD8B44E2CB4EC2 CRC64; |
| Alignment Scores: | |
| Pred. No.: | 3e-14 |
| Score: | 152.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 92.68% |
| DB: | 4 |
| US-09-938-703-3_COPY_792_885 (1-94) x Q9UN27 (1-339) | |
| QY | 2 TTAAGATGCTCATCTGGCGTGTCTGCCGCTGCTGTGTCATGG |
| Db | 189 LeuNlysleValIleLeuGlyLeuValLeuProLeuLeuValMetv |
| QY | 62 GGAATCCTAAAAACTCTGCTCGGTGTCGAAAT 94 |
| Db | 209 GlyIleLeuLysThrLeuLeuArgCysArgasn 219 |
| RESULT 6 | |
| Q9UN26 | PRELIMINARY; PRT; 339 AA. |
| ID | Q9UN26 |
| AC | Q9UN26 |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) |

SQ SEQUENCE 339 AA; 39146 MW; 10FE05FE5371D4B3 CRC64;

Alignment Scores:
Pred. No.: 3e-14 Length: 339
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9UN25 (1-339)

QY 2 TTAAGATAGTCATCTGGGCTGCTCCGCTGCTGTCATGTCATCTGCTACTCG 61

Db 189 LeuylsileValleLeuglyLeuValleProLeuLeuValMetValleCysTyrSer 208

QY 62 GGAATCCTAAAACTCTGCTCGTTCGGAAT 94

Db 209 GlylleLeuylsThrLeuLeuArgCysArgAsn 219

RESULT 8

Q9UN24
ID Q9UN24 PRELIMINARY; PRT; 339 AA.

AC Q9UN24;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).

GN CCR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.

RA Kunstan K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wollinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF161919; AAD47676.1; -

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF000001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; UNKNOWN_1.

DR PROSITE; PSS0262; G-PROTEIN_RECEP_FL_2; 1.

KW Receptor.

FT NON_TER 1 1

FT NON_TER 339 339

SQ SEQUENCE 339 AA; 39066 MW; 8BAF02E19423BF79 CRC64;

Alignment Scores:

Pred. No.: 3e-14 Length: 339
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9UN24 (1-339)

QY 2 TTAAGATAGTCATCTGGGCTGCTCCGCTGCTGTCATGTCATCTGCTACTCG 61

Db 189 LeuylsileValleLeuglyLeuValleProLeuLeuValMetValleCysTyrSer 208

QY 62 GGAATCCTAAAACTCTGCTCGTTCGGAAT 94

Db 209 GlylleLeuylsThrLeuLeuArgCysArgAsn 219

RESULT 9

Q9UN23

ID Q9UN23 PRELIMINARY; PRT; 339 AA.

AC Q9UN23;

DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).

GN CCR5.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.

RA Kunstan K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wollinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF161920; AAD47677.1; -

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF000001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; UNKNOWN_1.

DR PROSITE; PSS0262; G-PROTEIN_RECEP_FL_2; 1.

KW Receptor.

FT NON_TER 1 1

FT NON_TER 339 339

SQ SEQUENCE 339 AA; 39115 MW; 3C6369F922C91AA7 CRC64;

Alignment Scores:

Pred. No.: 3e-14 Length: 339
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x Q9UN23 (1-339)

QY 2 TTAAGATAGTCATCTGGGCTGCTCCGCTGCTGTCATGTCATCTGCTACTCG 61

Db 189 LeuylsileValleLeuglyLeuValleProLeuLeuValMetValleCysTyrSer 208

QY 62 GGAATCCTAAAACTCTGCTCGTTCGGAAT 94

Db 209 GlylleLeuylsThrLeuLeuArgCysArgAsn 219

RESULT 10

Q9UN29

ID Q9UN29 PRELIMINARY; PRT; 339 AA.

AC Q9UN29;

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE C-C chemokine receptor 5 (Fragment).

GN CCR5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.

RA Kunstan K., Chen Z., Korber B., Oprondek J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wollinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF161915; AAD47672.1; -

DR EMBL; AF161909; AAD47666.1; -

DR EMBL; AF161910; AAD47667.1; -

DR EMBL; AF161911; AAD47668.1; -

DR EMBL; AF161912; AAD47669.1; -

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF000001; 7tm_1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

QY 62 GGAATCCTAAAACTCTGCTTCGGTGTGCAAT 94

RP SEQUENCE FROM N.A.
RA kunstman K., Chen Z., Korber B., Stanton J., Agy M.
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF161886; AAD74645.1; -
InterPro: IPR000276; GPCR_Rhodopsn.

Job time : 7.73177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.p2p model

Run on: June 3, 2003, 18:39:54 ; Search time 1.52995 Seconds
(without alignments)
5096.611 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequence: 1 ATTAAGATAGTCACTGTGG.....CTCTGCTTCGGTGTGGAAT 94

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-O/cgm21/USPTO.spool/US09938703/runat_03062003_161406_22808/app_query.fasta_1.1870
-DB=SwissProt_40 -QFMT=Easton -SUFFIX=rs -MINMATCH=0.1 -LOOP=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703 -CGN_1_1_38 -runat_03062003_161406_22808 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARM_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 152 | 92.7 | 352 | 1 | CKR5_CERAE |
| 2 | 152 | 92.7 | 352 | 1 | CKR5_CERTO |
| 3 | 152 | 92.7 | 352 | 1 | CKR5_GORGO |
| 4 | 152 | 92.7 | 352 | 1 | CKR5_HUMAN |
| 5 | 152 | 92.7 | 352 | 1 | CKR5_HYLLE |
| 6 | 152 | 92.7 | 352 | 1 | CKR5_PANTR |
| 7 | 152 | 92.7 | 352 | 1 | CKR5_PAPHA |
| 8 | 152 | 92.7 | 352 | 1 | CKR5_PONPY |
| 9 | 152 | 92.7 | 352 | 1 | CKR5_PIGBI |
| 10 | 152 | 92.7 | 352 | 1 | CKR5_TRAPI |
| 11 | 152 | 92.7 | 352 | 1 | CKR5_TRAPI |
| 12 | 151 | 92.9 | 352 | 1 | CKR5_PYGNE |
| 13 | 149 | 90.9 | 352 | 1 | CKR5_NACMU |
| 14 | 134 | 81.7 | 360 | 1 | CKR2_MACMU |
| 15 | 134 | 81.7 | 374 | 1 | CKR2_HUMAN |
| 16 | 133 | 81.1 | 354 | 1 | CKR5_RAT |
| 17 | 131 | 79.9 | 354 | 1 | CKR5_MOUSE |
| 18 | 122 | 74.4 | 359 | 1 | CKR3_MOUSE |

| | | | | | |
|----|-----|------|-----|---|------------|
| 19 | 121 | 73.8 | 359 | 1 | CKR3_RAT |
| 20 | 118 | 72.0 | 373 | 1 | CKR2_MOUSE |
| 21 | 118 | 72.0 | 373 | 1 | CKR2_RAT |
| 22 | 113 | 68.9 | 358 | 1 | CKR3_CAVPO |
| 23 | 112 | 68.3 | 355 | 1 | CKR3_MACMU |
| 24 | 109 | 66.5 | 355 | 1 | CKR3_CERAE |
| 25 | 109 | 66.5 | 355 | 1 | CKR3_HUMAN |
| 26 | 106 | 64.6 | 355 | 1 | CKR1_HUMAN |
| 27 | 105 | 64.0 | 355 | 1 | CKR1_MACMU |
| 28 | 105 | 64.0 | 355 | 1 | CKR1_MOUSE |
| 29 | 95 | 57.9 | 360 | 1 | CKR4_HUMAN |
| 30 | 88 | 53.7 | 354 | 1 | C3X1_RAT |
| 31 | 84 | 51.2 | 354 | 1 | C3X1_MOUSE |
| 32 | 81 | 49.4 | 356 | 1 | CKR5_MOUSE |
| 33 | 80 | 48.8 | 342 | 1 | CKR6_CERAE |
| 34 | 80 | 48.8 | 342 | 1 | CKR6_HUMAN |
| 35 | 80 | 48.8 | 342 | 1 | CKR6_MACNE |
| 36 | 80 | 48.8 | 343 | 1 | CKR6_MACMU |
| 37 | 80 | 48.8 | 360 | 1 | CKR4_MOUSE |
| 38 | 79 | 48.2 | 269 | 1 | VC03_SPVKA |
| 39 | 79 | 48.2 | 355 | 1 | C3X1_HUMAN |
| 40 | 78 | 47.6 | 370 | 1 | VK02_SPVKA |
| 41 | 76 | 46.3 | 378 | 1 | CKR7_HUMAN |
| 42 | 75 | 45.7 | 355 | 1 | CKR8_HUMAN |
| 43 | 74 | 45.1 | 374 | 1 | CKR6_HUMAN |
| 44 | 72 | 43.9 | 356 | 1 | CKR8_MACMU |
| 45 | 71 | 43.3 | 378 | 1 | CKR7_MOUSE |

ALIGNMENTS

RESULT 1
CKR5_CERAE
ID CKR5_CERAE STANDARD; PRT; 352 AA.
AC P56493;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
GN CCR5 OR CCR5R5.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=98001387; PubMed=9343222;
RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
implicate specific amino acids in infections by simian and human
immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [2]

SEQUENCE FROM N.A.
RP Murayama Y., Matsunaga S., Inoue-Murayama M.;
RA "cDNA sequence of African green monkey CCR-5 chemokine receptor
gene.";
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

Alignment Scores:

RN [6] SEQUENCE FROM N.A., AND POLYMORPHISMS.
 RP MEDLINE-9802612; PubMed-9359654;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98049523; PubMed-9388201;
 RA Mumtaz S., Ahuja S.S., McDaniel B.L., Ahuja S.K.;
 RT "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts
 RT with 5'-end heterogeneity, dual promoter usage, and evidence for
 RT polymorphisms within the regulatory regions and noncoding exons.";
 RL J. Biol. Chem. 272:30662-30671(1997).
 RN [8]
 RP SEQUENCE FROM N.A., AND VARIANT ARG-178.
 RA Magierowska M., Barre-Sinoussi F., Issafras H., Theodorou I.,
 RA Debre P.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 RX MEDLINE-96260017; PubMed-8649511;
 RA Deng H., Liu R., Ellmeier W., Choe S., Unutmaz D., Burkhart M.,
 RA di Marzio P., Marmon S., Sutton R.E., Hill C.M., Davis C.B.,
 RA Peiper S.C., Schall T.J., Littman D.R., Landau N.R.;
 RT "Identification of a major co-receptor for primary isolates of
 RT HIV-1.";
 RL Nature 381:661-666(1996).
 RN [10]
 RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 RX MEDLINE-96260018; PubMed-8649512;
 RA Dragic T., Litwin V., Allaway G.P., Martin S.R., Huang Y.,
 RA Nagashima K.A., Cayanan C., Maddon P.J., Koup R.A., Moore J.P.,
 RA Paxton W.A.;
 RT "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor
 RT CC-CKR-5.";
 RL Nature 381:667-673(1996).
 RN [11]
 RP SULFATION.
 RX MEDLINE-99189752; PubMed-10089882;
 RA Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
 RA Gerard N.P., Garabed C., Sodroski J., Choe H.;
 RT "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
 RT entry.";
 RL Cell 96:667-676(1999).
 CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION. ACTS AS CO-RECEPTOR WITH CD4 FOR PRIMARY NON-
 CC SYNCTIUM-INDUCING STRAINS (NSI) (MACROPHAGE-TROPIC) OF HIV-1
 CC VIRUS. IT PROMOTES ENV-MEDIATED FUSION OF THE VIRUS.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: FOUND IN PROMYELOCYTIC CELLS.
 CC -!- PTM: SULFATION CONTRIBUTES TO THE EFFICIENCY OF HIV-1 ENTRY.
 CC -!- PTM: MODIFIED BY O-LINKED GLYCOSYLATION, BUT NOT BY N-LINKED
 CC GLYCOSYLATION.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X91492; CAAG2796.1; -;
 CC EMBL; U54994; AAC50598.1; -;
 CC EMBL; U57840; AAB17071.1; -;
 CC EMBL; U95626; AAB57793.1; -;
 CC EMBL; U83326; AAC51797.1; -;
 DR EMBL; AF011500; AAB65700.1; -;
 DR EMBL; AF011501; AAB65701.1; -;
 DR EMBL; AF011502; AAB65702.1; -;
 DR EMBL; AF011503; AAB65703.1; -;
 DR EMBL; AF011504; AAB65704.1; -;
 DR EMBL; AF011505; AAB65705.1; -;
 DR EMBL; AF011506; AAB65706.1; -;
 DR EMBL; AF011507; AAB65707.1; -;
 DR EMBL; AF011508; AAB65708.1; -;
 DR EMBL; AF011509; AAB65709.1; -;
 DR EMBL; AF011510; AAB65710.1; -;
 DR EMBL; AF011511; AAB65711.1; -;
 DR EMBL; AF011512; AAB65712.1; -;
 DR EMBL; AF011513; AAB65713.1; -;
 DR EMBL; AF011514; AAB65714.1; -;
 DR EMBL; AF011515; AAB65715.1; -;
 DR EMBL; AF011516; AAB65716.1; -;
 DR EMBL; AF011517; AAB65717.1; -;
 DR EMBL; AF011518; AAB65718.1; -;
 DR EMBL; AF011519; AAB65719.1; -;
 DR EMBL; AF011520; AAB65720.1; -;
 DR EMBL; AF011521; AAB65721.1; -;
 DR EMBL; AF011522; AAB65722.1; -;
 DR EMBL; AF011523; AAB65723.1; -;
 DR EMBL; AF011524; AAB65724.1; -;
 DR EMBL; AF011525; AAB65725.1; -;
 DR EMBL; AF011526; AAB65726.1; -;
 DR EMBL; AF011527; AAB65727.1; -;
 DR EMBL; AF011528; AAB65728.1; -;
 DR EMBL; AF011529; AAB65729.1; -;
 DR EMBL; AF011530; AAB65730.1; -;
 DR EMBL; AF011531; AAB65731.1; -;
 DR EMBL; AF011532; AAB65732.1; -;
 DR EMBL; AF011533; AAB65733.1; -;
 DR EMBL; AF011534; AAB65734.1; -;
 DR EMBL; AF011535; AAB65735.1; -;
 DR EMBL; AF011536; AAB65736.1; -;
 DR EMBL; AF011537; AAB65737.1; -;
 DR EMBL; AF011538; AAB65738.1; -;
 DR EMBL; AF011539; AAB65739.1; -;
 DR EMBL; AF011540; AAB65740.1; -;
 DR EMBL; AF011541; AAB65741.1; -;
 DR EMBL; AF011542; AAB65742.1; -;
 DR EMBL; AF011543; AAB65743.1; -;
 DR EMBL; AF011544; AAB65744.1; -;
 DR EMBL; AF011545; AAB65745.1; -;
 DR EMBL; AF011546; AAB65746.1; -;
 DR EMBL; AF011547; AAB65747.1; -;
 DR EMBL; AF011548; AAB65748.1; -;
 DR EMBL; AF011549; AAB65749.1; -;
 DR EMBL; AF011550; AAB65750.1; -;
 DR EMBL; AF011551; AAB65751.1; -;
 DR EMBL; AF011552; AAB65752.1; -;
 DR EMBL; AF011553; AAB65753.1; -;
 DR EMBL; AF011554; AAB65754.1; -;
 DR EMBL; AF011555; AAB65755.1; -;
 DR EMBL; AF011556; AAB65756.1; -;
 DR EMBL; AF011557; AAB65757.1; -;
 DR EMBL; AF011558; AAB65758.1; -;
 DR EMBL; AF011559; AAB65759.1; -;
 DR EMBL; AF011560; AAB65760.1; -;
 DR EMBL; AF011561; AAB65761.1; -;
 DR EMBL; AF011562; AAB65762.1; -;
 DR EMBL; AF011563; AAB65763.1; -;
 DR EMBL; AF011564; AAB65764.1; -;
 DR EMBL; AF011565; AAB65765.1; -;
 DR EMBL; AF011566; AAB65766.1; -;
 DR EMBL; AF011567; AAB65767.1; -;
 DR EMBL; AF011568; AAB65768.1; -;
 DR EMBL; AF011569; AAB65769.1; -;
 DR EMBL; AF011570; AAB65770.1; -;
 DR EMBL; AF011571; AAB65771.1; -;
 DR EMBL; AF011572; AAB65772.1; -;
 DR EMBL; AF011573; AAB65773.1; -;
 DR EMBL; AF011574; AAB65774.1; -;
 DR EMBL; AF011575; AAB65775.1; -;
 DR EMBL; AF011576; AAB65776.1; -;
 DR EMBL; AF011577; AAB65777.1; -;
 DR EMBL; AF011578; AAB65778.1; -;
 DR EMBL; AF011579; AAB65779.1; -;
 DR EMBL; AF011580; AAB65780.1; -;
 DR EMBL; AF011581; AAB65781.1; -;
 DR EMBL; AF011582; AAB65782.1; -;
 DR EMBL; AF011583; AAB65783.1; -;
 DR EMBL; AF011584; AAB65784.1; -;
 DR EMBL; AF011585; AAB65785.1; -;
 DR EMBL; AF011586; AAB65786.1; -;
 DR EMBL; AF011587; AAB65787.1; -;
 DR EMBL; AF011588; AAB65788.1; -;
 DR EMBL; AF011589; AAB65789.1; -;
 DR EMBL; AF011590; AAB65790.1; -;
 DR EMBL; AF011591; AAB65791.1; -;
 DR EMBL; AF011592; AAB65792.1; -;
 DR EMBL; AF011593; AAB65793.1; -;
 DR EMBL; AF011594; AAB65794.1; -;
 DR EMBL; AF011595; AAB65795.1; -;
 DR EMBL; AF011596; AAB65796.1; -;
 DR EMBL; AF011597; AAB65797.1; -;
 DR EMBL; AF011598; AAB65798.1; -;
 DR EMBL; AF011599; AAB65799.1; -;
 DR EMBL; AF011600; AAB65800.1; -;
 DR EMBL; AF011601; AAB65801.1; -;
 DR EMBL; AF011602; AAB65802.1; -;
 DR EMBL; AF011603; AAB65803.1; -;
 DR EMBL; AF011604; AAB65804.1; -;
 DR EMBL; AF011605; AAB65805.1; -;
 DR EMBL; AF011606; AAB65806.1; -;
 DR EMBL; AF011607; AAB65807.1; -;
 DR EMBL; AF011608; AAB65808.1; -;
 DR EMBL; AF011609; AAB65809.1; -;
 DR EMBL; AF011610; AAB65810.1; -;
 DR EMBL; AF011611; AAB65811.1; -;
 DR EMBL; AF011612; AAB65812.1; -;
 DR EMBL; AF011613; AAB65813.1; -;
 DR EMBL; AF011614; AAB65814.1; -;
 DR EMBL; AF011615; AAB65815.1; -;
 DR EMBL; AF011616; AAB65816.1; -;
 DR EMBL; AF011617; AAB65817.1; -;
 DR EMBL; AF011618; AAB65818.1; -;
 DR EMBL; AF011619; AAB65819.1; -;
 DR EMBL; AF011620; AAB65820.1; -;
 DR EMBL; AF011621; AAB65821.1; -;
 DR EMBL; AF011622; AAB65822.1; -;
 DR EMBL; AF011623; AAB65823.1; -;
 DR EMBL; AF011624; AAB65824.1; -;
 DR EMBL; AF011625; AAB65825.1; -;
 DR EMBL; AF011626; AAB65826.1; -;
 DR EMBL; AF011627; AAB65827.1; -;
 DR EMBL; AF011628; AAB65828.1; -;
 DR EMBL; AF011629; AAB65829.1; -;
 DR EMBL; AF011630; AAB65830.1; -;
 DR EMBL; AF011631; AAB65831.1; -;
 DR EMBL; AF011632; AAB65832.1; -;
 DR EMBL; AF011633; AAB65833.1; -;
 DR EMBL; AF011634; AAB65834.1; -;
 DR EMBL; AF011635; AAB65835.1; -;
 DR EMBL; AF011636; AAB65836.1; -;
 DR EMBL; AF011637; AAB65837.1; -;
 DR EMBL; AF011638; AAB65838.1; -;
 DR EMBL; AF011639; AAB65839.1; -;
 DR EMBL; AF011640; AAB65840.1; -;
 DR EMBL; AF011641; AAB65841.1; -;
 DR EMBL; AF011642; AAB65842.1; -;
 DR EMBL; AF011643; AAB65843.1; -;
 DR EMBL; AF011644; AAB65844.1; -;
 DR EMBL; AF011645; AAB65845.1; -;
 DR EMBL; AF011646; AAB65846.1; -;
 DR EMBL; AF011647; AAB65847.1; -;
 DR EMBL; AF011648; AAB65848.1; -;
 DR EMBL; AF011649; AAB65849.1; -;
 DR EMBL; AF011650; AAB65850.1; -;
 DR EMBL; AF011651; AAB65851.1; -;
 DR EMBL; AF011652; AAB65852.1; -;
 DR EMBL; AF011653; AAB65853.1; -;
 DR EMBL; AF011654; AAB65854.1; -;
 DR EMBL; AF011655; AAB65855.1; -;
 DR EMBL; AF011656; AAB65856.1; -;
 DR EMBL; AF011657; AAB65857.1; -;
 DR EMBL; AF011658; AAB65858.1; -;
 DR EMBL; AF011659; AAB65859.1; -;
 DR EMBL; AF011660; AAB65860.1; -;
 DR EMBL; AF011661; AAB65861.1; -;
 DR EMBL; AF011662; AAB65862.1; -;
 DR EMBL; AF011663; AAB65863.1; -;
 DR EMBL; AF011664; AAB65864.1; -;
 DR EMBL; AF011665; AAB65865.1; -;
 DR EMBL; AF011666; AAB65866.1; -;
 DR EMBL; AF011667; AAB65867.1; -;
 DR EMBL; AF011668; AAB65868.1; -;
 DR EMBL; AF011669; AAB65869.1; -;
 DR EMBL; AF011670; AAB65870.1; -;
 DR EMBL; AF011671; AAB65871.1; -;
 DR EMBL; AF011672; AAB65872.1; -;
 DR EMBL; AF011673; AAB65873.1; -;
 DR EMBL; AF011674; AAB65874.1; -;
 DR EMBL; AF011675; AAB65875.1; -;
 DR EMBL; AF011676; AAB65876.1; -;
 DR EMBL; AF011677; AAB65877.1; -;
 DR EMBL; AF011678; AAB65878.1; -;
 DR EMBL; AF011679; AAB65879.1; -;
 DR EMBL; AF011680; AAB65880.1; -;
 DR EMBL; AF011681; AAB65881.1; -;
 DR EMBL; AF011682; AAB65882.1; -;
 DR EMBL; AF011683; AAB65883.1; -;
 DR EMBL; AF011684; AAB65884.1; -;
 DR EMBL; AF011685; AAB65885.1; -;
 DR EMBL; AF011686; AAB65886.1; -;
 DR EMBL; AF011687; AAB65887.1; -;
 DR EMBL; AF011688; AAB65888.1; -;
 DR EMBL; AF011689; AAB65889.1; -;
 DR EMBL; AF011690; AAB65890.1; -;
 DR EMBL; AF011691; AAB65891.1; -;
 DR EMBL; AF011692; AAB65892.1; -;
 DR EMBL; AF011693; AAB65893.1; -;
 DR EMBL; AF011694; AAB65894.1; -;
 DR EMBL; AF011695; AAB65895.1; -;
 DR EMBL; AF011696; AAB65896.1; -;
 DR EMBL; AF011697; AAB65897.1; -;
 DR EMBL; AF011698; AAB65898.1; -;
 DR EMBL; AF011699; AAB65899.1; -;
 DR EMBL; AF011700; AAB65900.1; -;
 DR EMBL; AF011701; AAB65901.1; -;
 DR EMBL; AF011702; AAB65902.1; -;
 DR EMBL; AF011703; AAB65903.1; -;
 DR EMBL; AF011704; AAB65904.1; -;
 DR EMBL; AF011705; AAB65905.1; -;
 DR EMBL; AF011706; AAB65906.1; -;
 DR EMBL; AF011707; AAB65907.1; -;
 DR EMBL; AF011708; AAB65908.1; -;
 DR EMBL; AF011709; AAB65909.1; -;
 DR EMBL; AF011710; AAB65910.1; -;
 DR EMBL; AF011711; AAB65911.1; -;
 DR EMBL; AF011712; AAB65912.1; -;
 DR EMBL; AF011713; AAB65913.1; -;
 DR EMBL; AF011714; AAB65914.1; -;
 DR EMBL; AF011715; AAB65915.1; -;
 DR EMBL; AF011716; AAB65916.1; -;
 DR EMBL; AF011717; AAB65917.1; -;
 DR EMBL; AF011718; AAB65918.1; -;
 DR EMBL; AF011719; AAB65919.1; -;
 DR EMBL; AF011720; AAB65920.1; -;
 DR EMBL; AF011721; AAB65921.1; -;
 DR EMBL; AF011722; AAB65922.1; -;
 DR EMBL; AF011723; AAB65923.1; -;
 DR EMBL; AF011724; AAB65924.1; -;
 DR EMBL; AF011725; AAB65925.1; -;
 DR EMBL; AF011726; AAB65926.1; -;
 DR EMBL; AF011727; AAB65927.1; -;
 DR EMBL; AF011728; AAB65928.1; -;
 DR EMBL; AF011729; AAB65929.1; -;
 DR EMBL; AF011730; AAB65930.1; -;
 DR EMBL; AF011731; AAB65931.1; -;
 DR EMBL; AF011732; AAB65932.1; -;
 DR EMBL; AF011733; AAB65933.1; -;
 DR EMBL; AF011734; AAB65934.1; -;
 DR EMBL; AF011735; AAB65935.1; -;
 DR EMBL; AF011736; AAB65936.1; -;
 DR EMBL; AF011737; AAB65937.1; -;
 DR EMBL; AF011738; AAB65938.1; -;
 DR EMBL; AF011739; AAB65939.1; -;
 DR EMBL; AF011740; AAB65940.1; -;
 DR EMBL; AF011741; AAB65941.1; -;
 DR EMBL; AF011742; AAB65942.1; -;
 DR EMBL; AF011743; AAB65943.1; -;
 DR EMBL; AF011744; AAB65944.1; -;
 DR EMBL; AF011745; AAB65945.1; -;
 DR EMBL; AF011746; AAB65946.1; -;
 DR EMBL; AF011747; AAB65947.1; -;
 DR EMBL; AF011748; AAB65948.1; -;
 DR EMBL; AF011749; AAB65949.1; -;
 DR EMBL; AF011750; AAB65950.1; -;
 DR EMBL; AF011751; AAB65951.1; -;
 DR EMBL; AF011752; AAB65952.1; -;
 DR EMBL; AF011753; AAB65953.1; -;
 DR EMBL; AF011754; AAB65954.1; -;
 DR EMBL; AF011755; AAB65955.1; -;
 DR EMBL; AF011756; AAB65956.1; -;
 DR EMBL; AF011757; AAB65957.1; -;
 DR EMBL; AF011758; AAB65958.1; -;
 DR EMBL; AF011759; AAB65959.1; -;
 DR EMBL; AF011760; AAB65960.1; -;
 DR EMBL; AF011761; AAB65961.1; -;
 DR EMBL; AF011762; AAB65962.1; -;
 DR EMBL; AF011763; AAB65963.1; -;
 DR EMBL; AF011764; AAB65964.1; -;
 DR EMBL; AF011765; AAB65965.1; -;
 DR EMBL; AF011766; AAB65966.1; -;
 DR EMBL; AF011767; AAB65967.1; -;
 DR EMBL; AF011768; AAB65968.1; -;
 DR EMBL; AF011769; AAB65969.1; -;
 DR EMBL; AF011770; AAB65970.1; -;
 DR EMBL; AF011771; AAB65971.1; -;
 DR EMBL; AF011772; AAB65972.1; -;
 DR EMBL; AF011773; AAB65973.1; -;
 DR EMBL; AF011774; AAB65974.1; -;
 DR EMBL; AF011775; AAB65975.1; -;
 DR EMBL; AF011776; AAB65976.1; -;
 DR EMBL; AF011777; AAB65977.1; -;
 DR EMBL; AF011778; AAB65978.1; -;
 DR EMBL; AF011779; AAB65979.1; -;
 DR EMBL; AF011780; AAB65980.1; -;
 DR EMBL; AF011781; AAB65981.1; -;
 DR EMBL; AF011782; AAB65982.1; -;
 DR EMBL; AF011783; AAB65983.1; -;
 DR EMBL; AF011784; AAB65984.1; -;
 DR EMBL; AF011785; AAB65985.1; -;
 DR EMBL; AF011786; AAB65986.1; -;
 DR EMBL; AF011787; AAB65987.1; -;
 DR EMBL; AF011788; AAB65988.1; -;
 DR EMBL; AF011789; AAB65989.1; -;
 DR EMBL; AF011790; AAB65990.1; -;
 DR EMBL; AF011791; AAB65991.1; -;
 DR EMBL; AF011792; AAB65992.1; -;
 DR EMBL; AF011793; AAB65993.1; -;
 DR EMBL; AF011794; AAB65994.1; -;
 DR EMBL; AF011795; AAB65995.1; -;
 DR EMBL; AF011796; AAB65996.1; -;
 DR EMBL; AF011797; AAB65997.1; -;
 DR EMBL; AF011798; AAB65998.1; -;
 DR EMBL; AF011799; AAB65999.1; -;
 DR EMBL; AF011800; AAB66000.1; -;
 DR EMBL; AF011801; AAB66001.1; -;
 DR EMBL; AF011802; AAB66002.1; -;
 DR EMBL; AF011803; AAB66003.1; -;
 DR EMBL; AF011804; AAB66004.1; -;
 DR EMBL; AF011805; AAB66005.1; -;
 DR EMBL; AF011806; AAB66006.1; -;
 DR EMBL; AF011807; AAB66007.1; -;
 DR EMBL; AF011808; AAB66008.1; -;
 DR EMBL; AF011809; AAB66009.1; -;
 DR EMBL; AF011810; AAB66010.1; -;
 DR EMBL; AF011811; AAB66011.1; -;
 DR EMBL; AF011812; AAB66012.1; -;
 DR EMBL; AF011813; AAB66013.1; -;
 DR EMBL; AF011814; AAB66014.1; -;
 DR EMBL; AF011815; AAB66015.1; -;
 DR EMBL; AF011816; AAB66016.1; -;
 DR EMBL; AF011817; AAB66017.1; -;
 DR EMBL; AF011818; AAB66018.1; -;
 DR EMBL; AF011819; AAB66019.1; -;
 DR EMBL; AF011820; AAB66020.1; -;
 DR EMBL; AF011821; AAB66021.1; -;
 DR EMBL; AF011822; AAB66022.1; -;
 DR EMBL; AF011823; AAB66023.1; -;
 DR EMBL; AF011824; AAB66024.1; -;
 DR EMBL; AF011825; AAB66025.1; -;
 DR EMBL; AF011826; AAB66026.1; -;
 DR EMBL; AF011827; AAB66027.1; -;
 DR EMBL; AF011828; AAB66028.1; -;
 DR EMBL; AF011829; AAB66029.1; -;
 DR EMBL; AF011830; AAB66030.1; -;
 DR EMBL; AF011831; AAB66031.1; -;
 DR EMBL; AF011832; AAB66032.1; -;
 DR EMBL; AF011833; AAB66033.1; -;
 DR EMBL; AF011834; AAB66034.1; -;
 DR EMBL; AF011835; AAB66035.1; -;
 DR EMBL; AF011836; AAB66036.1; -;
 DR EMBL; AF011837; AAB66037.1; -;
 DR EMBL; AF011838; AAB66038.1; -;
 DR EMBL; AF011839; AAB66039.1; -;
 DR EMBL; AF011840; AAB66040.1; -;
 DR EMBL; AF011841; AAB66041.1; -;
 DR EMBL; AF011842; AAB66042.1; -;
 DR EMBL; AF011843; AAB66043.1; -;
 DR EMBL; AF011844; AAB66044.1; -;
 DR EMBL; AF011845; AAB66045.1; -;
 DR EMBL; AF011846; AAB66046.1; -;
 DR EMBL; AF011847; AAB66047.1; -;
 DR EMBL; AF011848; AAB66048.1; -;
 DR EMBL; AF011849; AAB66049.1; -;
 DR EMBL; AF011850; AAB66050.1; -;
 DR EMBL; AF011851; AAB66051.1; -;
 DR EMBL; AF011852; AAB66052.1; -;
 DR EMBL; AF011853; AAB66053.1; -;
 DR EMBL; AF011854; AAB66054.1; -;
 DR EMBL; AF011855; AAB66055.1; -;
 DR EMBL; AF011856; AAB66056.1; -;
 DR EMBL; AF011857; AAB66057.1; -;
 DR EMBL; AF011858; AAB66058.1; -;
 DR EMBL; AF011859; AAB66059.1; -;
 DR EMBL; AF011860; AAB66060.1; -;
 DR EMBL; AF011861; AAB66061.1; -;
 DR EMBL; AF011862; AAB66062.1; -;
 DR EMBL; AF011863; AAB66063.1; -;
 DR EMBL; AF011864; AAB66064.1; -;
 DR EMBL; AF011865; AAB66065.1; -;
 DR EMBL; AF011866; AAB6

| | | | |
|----|-----|----------|-----|
| FT | 236 | TRANSMEM | 260 |
| FT | 219 | DOMAIN | 235 |

RT Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.¹⁰

| | | | | |
|----|----------|---------|-----------|----------------------------|
| FT | DOMAIN | 1 | 30 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 31 | 58 | 1 (POTENTIAL). |
| FT | DOMAIN | 59 | 68 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 69 | 89 | 2 (POTENTIAL). |
| FT | DOMAIN | 90 | 102 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 103 | 124 | 3 (POTENTIAL). |
| FT | DOMAIN | 125 | 141 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 142 | 166 | 4 (POTENTIAL). |
| FT | DOMAIN | 167 | 198 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 199 | 218 | 5 (POTENTIAL). |
| FT | DOMAIN | 219 | 235 | CYTOPLASMIC (POTENTIAL). |
| FT | TRANSMEM | 236 | 260 | 6 (POTENTIAL). |
| FT | DOMAIN | 261 | 277 | EXTRACELLULAR (POTENTIAL). |
| FT | TRANSMEM | 278 | 301 | 7 (POTENTIAL). |
| FT | DOMAIN | 302 | 352 | CYTOPLASMIC (POTENTIAL). |
| FT | DISULFID | 101 | 178 | BY SIMILARITY. |
| FT | MOD_RES | 3 | 3 | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 10 | 10 | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 14 | 14 | SULFATION (BY SIMILARITY). |
| FT | MOD_RES | 15 | 15 | SULFATION (BY SIMILARITY). |
| SQ | SEQUENCE | 352 AA; | 40527 MW; | F4E2P47135AF58A CRC64; |

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5,57e-14 | Length: | 352 |
| Score: | 152.00 | Matches: | 31 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 92.68% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-09-938-703-3_COPY_792.885 (1-94) x CKR5_PONPY (1-352)

Qy 2 TAAAGATAGTCATCTTGCGGCTGCTGCTCCGCTGCTGTCATGCTCATCTGCTACTCG 61
 |||
 Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215
 |||

Qy 62 GGAATCCTAAAACTCTGCTCGGTGTCGGAAT 94
 |||
 Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
 |||

RESULT 9

| | | | | |
|----|--|-----------|------|---------|
| ID | CKR5_PYGBI | STANDARD; | PRT; | 352 AA. |
| AC | 097880; | | | |
| DT | 30-MAY-2000 (Rel. 39, Created) | | | |
| DT | 30-MAY-2000 (Rel. 39, Last annotation update) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DE | C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5). | | | |
| GN | CCR5 OR CMKR5 | | | |
| OS | Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti); | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae; | | | |
| OC | Pygathrix. | | | |
| OX | NCBI_TaxID=61621; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=99416438; PubMed=10486970; | | | |
| RA | Zhang Y.-W., Ryder O.A., Zhang Y.-P.; | | | |
| RT | "sequence evolution of the CCR5 chemokine receptor gene in primates."; | | | |
| RL | Mol. Biol. Evol. 16:1145-1154(1999). | | | |
| CC | -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA, | | | |
| CC | MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY | | | |
| CC | INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE | | | |
| CC | IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR | | | |
| CC | DIFFERENTIATION. | | | |
| CC | SUBCELLULAR LOCATION: Integral membrane protein. | | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
DIFFERENTIATION.

- - SUBCELLULAR LOCATION: Integral membrane protein.

- - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AF075445; AAD19857.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN
CC PROSITE; PS00237; G-PROTEIN_RECEP_FL1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECEP_FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
CC DOMAIN 1 30
CC TRANSMEM 31 58 1 (POTENTIAL).
FT

| | | | |
|----------|---------|-----------|----------------------------|
| TRANSEM | 69 | 89 | 2 (POTENTIAL). |
| DOMAIN | 90 | 102 | EXTRACELLULAR (POTENTIAL). |
| TRANSEM | 103 | 124 | 3 (POTENTIAL). |
| DOMAIN | 125 | 141 | CYTOPLASMIC (POTENTIAL). |
| TRANSEM | 142 | 166 | 4 (POTENTIAL). |
| DOMAIN | 167 | 198 | EXTRACELLULAR (POTENTIAL). |
| TRANSEM | 199 | 218 | 5 (POTENTIAL). |
| DOMAIN | 219 | 235 | CYTOPLASMIC (POTENTIAL). |
| TRANSEM | 236 | 260 | 6 (POTENTIAL). |
| DOMAIN | 261 | 277 | EXTRACELLULAR (POTENTIAL). |
| TRANSEM | 278 | 301 | 7 (POTENTIAL). |
| DOMAIN | 302 | 352 | CYTOPLASMIC (POTENTIAL). |
| DISULFID | 101 | 178 | BY SIMILARITY. |
| MOD_RES | 3 | 3 | SULFATION (BY SIMILARITY). |
| MOD_RES | 10 | 10 | SULFATION (BY SIMILARITY). |
| MOD_RES | 14 | 14 | SULFATION (BY SIMILARITY). |
| MOD_RES | 15 | 15 | SULFATION (BY SIMILARITY). |
| SEQUENCE | 352 AA; | 40585 MW; | 43666F142730F938F CRC64. |

[illegible]

| | |
|------------------------|-------------------------------------|
| Alignment Scores: | |
| Pred. No.: | 5, 57e-14 |
| Score: | 152.00 |
| Length: | 352 |
| Matches: | 31 |
| Percent Similarity: | 100.00% |
| Conservative: | 0 |
| Best Local Similarity: | 100.00% |
| Mismatches: | 0 |
| Query Match: | 92.68% |
| Indels: | 0 |
| Gaps: | 0 |
| DB: | 1 |
| US-09-938-703-3 COPY | 792 885 (1-94) x CKR5_PYGIB (1-352) |

Search completed: June 3, 2003, 18:46:03
Job time : 3.52995 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_p2p model

Run on: June 3, 2003, 18:42:09 ; Search time 3.30469 Seconds
(without alignments)
5468.980 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164
Sequence: 1 ATTAAGATAGTATCATCTGG.....CTCTGCTTCGGTGCAGAAAT 94

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delep 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q/cgr2_1/USPTO_spool/US9938703/runat_03062003_161407_22855/app_query.fasta_1.1870
-DB=PIR_73 -QFWT=Eastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9938703 -ECGN_1_1_89 -runat_03062003_161407_22855 -NCP0=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PIR_73:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | ID | Description |
|------------|-------------|--------|-----|-------------|
| 1 | 152 | 92.7 | 352 | 2 A43113 |
| 2 | 134 | 81.7 | 360 | 2 JC2443 |
| 3 | 134 | 81.7 | 374 | 2 I38450 |
| 4 | 122 | 74.4 | 359 | 2 I49341 |
| 5 | 109 | 66.5 | 355 | 2 G02436 |
| 6 | 106 | 64.6 | 355 | 2 A45177 |
| 7 | 105 | 64.0 | 355 | 2 I49339 |
| 8 | 95 | 57.9 | 350 | 2 A57160 |
| 9 | 88 | 53.7 | 354 | 2 I58186 |
| 10 | 84 | 51.2 | 383 | 2 S55594 |
| 11 | 81 | 49.4 | 356 | 2 I49340 |
| 12 | 80 | 48.8 | 360 | 2 JC4587 |
| 13 | 79 | 48.2 | 355 | 2 JC4304 |
| 14 | 76 | 46.3 | 378 | 2 A45680 |

| | | | | |
|----|----|------|-----|----------|
| 15 | 76 | 45.3 | 378 | 2 B55735 |
| 16 | 75 | 45.7 | 355 | 2 JC5067 |
| 17 | 74 | 45.1 | 369 | 2 JC5068 |
| 18 | 71 | 43.3 | 378 | 2 A55735 |
| 19 | 68 | 41.5 | 367 | 2 JB0349 |
| 20 | 66 | 40.2 | 350 | 2 JN0621 |
| 21 | 66 | 40.2 | 354 | 2 B55733 |
| 22 | 66 | 40.2 | 354 | 2 A3669 |
| 23 | 66 | 40.2 | 355 | 2 JQ1231 |
| 24 | 66 | 40.2 | 363 | 2 I48261 |
| 25 | 66 | 40.2 | 363 | 2 A49092 |
| 26 | 66 | 40.2 | 363 | 2 JC2543 |
| 27 | 64 | 39.0 | 356 | 2 S42096 |
| 28 | 64 | 39.0 | 358 | 2 A5752 |
| 29 | 63 | 38.4 | 359 | 2 A48921 |
| 30 | 63 | 38.4 | 360 | 2 A53611 |
| 31 | 63 | 38.4 | 361 | 2 B45680 |
| 32 | 62 | 37.8 | 374 | 2 S32785 |
| 33 | 60 | 36.6 | 350 | 2 A39445 |
| 34 | 60 | 36.6 | 375 | 2 JC5509 |
| 35 | 59 | 36.0 | 344 | 2 T09508 |
| 36 | 59 | 36.0 | 358 | 2 G02670 |
| 37 | 59 | 36.0 | 374 | 2 S42628 |
| 38 | 59 | 36.0 | 375 | 2 JC5069 |
| 39 | 57 | 34.8 | 314 | 2 T46279 |
| 40 | 57 | 34.8 | 366 | 2 I48469 |
| 41 | 57 | 34.8 | 370 | 2 JC5549 |
| 42 | 57 | 34.8 | 482 | 2 S65766 |
| 43 | 56 | 34.1 | 240 | 2 S55480 |
| 44 | 56 | 34.1 | 350 | 1 A37963 |
| 45 | 56 | 34.1 | 380 | 2 JC2434 |

ALIGNMENTS

RESULT 1

A43113

Chemokine (C-C) receptor 5 - human

N:Alternate names: C-CR-5; CCR5

C:Species: Homo sapiens (man)

C>Date: 12-Jul-1996 #sequence,revision 12-Jul-1996 #text_change 20-Jun-2000

C:Accession: A43113; S71808; A58834; A58832; G02653; A58833

R:Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.

Biochemistry 35, 3362-3367, 1996

A>Title: Molecular cloning and functional expression of a new human CC-chemokine rece

A:Reference number: A43113; MUID:96241590; PMID:8639485

A:Accession: A43113

A:Molecule type: mRNA

A:Residues: 1-352 <SAMI>

A:Cross-references: GB:X01492; NID:gl262810; PIDN:CAA62796.1; PID:gl262811

R:Samson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Lienard, C.; Farber, C.M.; Sarag

M.; Imal, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;

Nature 382, 722-725, 1996

A>Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele

A:Reference number: S71808; MUID:96345670; PMID:8751444

A:Accession: S71808

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 182-206;207-230 <SAM2>

A:Accession: A58834

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-184, 'IKDSHLCAGPAAACHGHLILGNPKNSAVSK' <SAM3>

A:Cross-references: GB:X99393; NID:gl524062; PIDN:CAA67767.1; PID:gl524063

A>Note: This frameshift mutation results in a non-functional receptor but confers a c

nd may have had a selective advantage by conferring resistance to Versinia plague inf

J. Combadiere, C.; Aluja, S.K.; Tiffany, H.L.; Murphy, P.M.

T. Leukoc. Biol. 60, 147-152, 1996

A>Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine

A:Reference number: A58832; MUID:96295970; PMID:8699119

A:Accession: A58832

A:Molecule type: mRNA

A:Residues: 1-352 <COMI>

A;Cross-references: DDBJ:D29984; NID:9531246; PIDN:BR006253.1; PID:9531247
R;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994

A;Title: Molecular cloning and functional expression of two monocyte chemoattractant
A;Reference number: A53477; MUID:94195821; PMID:8146186

A;Accession: I38463
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-360 <RES>
A;Cross-references: EMBL:U03905; NID:9472557; PIDN:AAA19120.1; PID:9472558
C;Genetics:
A;Gene: GDB:CMKBR2

A;Cross-references: GDB:337364; OMIM:601267
A;Map position: 3p21-3p21
C;Superfamily: vertebrate rhodopsin

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane
F;81-100/Domain: transmembrane #status predicted <TM2>
F;81-100/Domain: transmembrane #status predicted <TM2>
F;115-136/Domain: transmembrane #status predicted <TM2>
F;154-178/Domain: transmembrane #status predicted <TM4>
F;207-226/Domain: transmembrane #status predicted <TM5>
F;244-268/Domain: transmembrane #status predicted <TM6>
F;287-309/Domain: transmembrane #status predicted <TM7>
F;14/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;113-190/pisulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 5,5e-11 Length: 360
Score: 134.00 Matches: 26
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 96.30% Mismatches: 0
Query Match: 81.71% Indels: 0
DB: 2 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x JC2443 (1-360)

Qy 14 ATCTGGGGCTGGTCTGGCGCTGCTGTGTCATGCTACTGCTACTGCGGAATCCTAAAA 73
|||||
Db 208 ILeucluyLeuValleuProleuLeuMetValleuIleCysTyrSerGlyLeuLeuLys 227
|||||

Qy 74 ACTCTGCTCTGGTGTGGAAT 94
|||||
Db 228 ThrLeuLeuArgCysAlaGln 234
|||||

RESULT 3
138450
chemokine (C-C) receptor 2, splice form A - human
N;Alternate names: C-C CKR-2; monocyte chemoattractant protein 1 receptor; monocyte
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 13-Aug-1999
C;Accession: I38450
R;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994

A;Title: Molecular cloning and functional expression of two monocyte chemoattractant
A;Reference number: A53477; MUID:94195821; PMID:8146186

A;Accession: I38450
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-374 <RES>
A;Cross-references: EMBL:U03882; NID:9472555; PIDN:AAA19119.1; PID:9472556
C;Genetics:
A;Gene: GDB:CMKBR2

A;Cross-references: GDB:337364; OMIM:601267
A;Map position: 3p21-3p21
C;Superfamily: vertebrate rhodopsin

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane
F;44-68/Domain: transmembrane #status predicted <TM1>
F;79-99/Domain: transmembrane #status predicted <TM2>
F;115-136/Domain: transmembrane #status predicted <TM3>
F;154-178/Domain: transmembrane #status predicted <TM4>
F;208-226/Domain: transmembrane #status predicted <TM5>
F;244-265/Domain: transmembrane #status predicted <TM6>
F;292-309/Domain: transmembrane #status predicted <TM7>

R:Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogewerf, A.J.; Proudfoot, A.E.I.
J. Biol. Chem. 270, 19495-19500, 1995

A:Title: Molecular cloning and functional expression of a novel CC chemokine receptor
A:Reference number: A57160; MUID:95370289; PMID:7642634
A:Accession: A57160

A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-360 <P>
A:Cross-references: GB:X85740; NID:g1370103; PIDN:CAA59743.1; PID:g971452
A>Note: source clone K5-5
C:Genetics:
A:Gene: GDB:CMKBR4
A:Cross-references: GDB:677463
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:40-65/Domain: transmembrane #status predicted <TM1>
F:76-97/Domain: transmembrane #status predicted <TM2>
F:112-133/Domain: transmembrane #status predicted <TM3>
F:151-175/Domain: transmembrane #status predicted <TM4>
F:208-226/Domain: transmembrane #status predicted <TM5>
F:243-264/Domain: transmembrane #status predicted <TM6>
F:291-308/Domain: transmembrane #status predicted <TM7>
F:29-276,110-187/Disulfide bonds: #status predicted
F:72,350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predi
F:145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicte
F:183,194/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte

Alignment Scores:

| Pred. No.: | 2,26e-05 | Length: | 360 |
|------------------------|----------|---------------|-----|
| Score: | 95.00% | Matches: | 18 |
| Percent Similarity: | 95.00% | Conservative: | 7 |
| Best Local Similarity: | 58.06% | Mismatches: | 6 |
| Query Match: | 57.93% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-938-703-3_COPY_792_885 (1-94) x A57160 (1-360)

Qy 2 TTAAGATGATCATCTGGCGTGTGCCTCCGTCGTGTGCTACTGCTACTCG 61
|||||
Db 204 LeuGLuIleAAnIIetLeuGLyLeuValIleProLeuGLyIleMetLeuPhcYtYrSer 223
|||
Qy 62 GGAATCCTAAACATCTGCTCGTGTCGGAAT 94
|||||
Db 224 MetIleIleArgThrLeuGLnHisCysLysAsn 234
|||

RESULT 9
158186

Probable G protein-coupled receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 21-Jul-2000
C:Accession: I58186
R:Harrison, J.K.; Barber, C.M.; Lynch, K.R.
Neurosci. Lett. 169, 85-89, 1994

A:Title: cDNA cloning of a G-protein-coupled receptor expressed in rat spinal cord an
A:Reference number: I58186; MUID:94323113; PMID:8047298
A:Accession: I58186
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-354 <RES>
A:Cross-references: EMBL:U04808; NID:g2556635; PIDN:AAB87093.1; PID:g439861
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor

Alignment Scores:

| Pred. No.: | 0.00023 | Length: | 354 |
|------------------------|---------|---------------|-----|
| Score: | 88.00 | Matches: | 16 |
| Percent Similarity: | 73.33% | Conservative: | 6 |
| Best Local Similarity: | 53.33% | Mismatches: | 8 |
| Query Match: | 53.66% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-09-938-703-3 COPY 792 885 (1-94) x I49340 (1-356)

| Result No. | Query | | | DB | ID | Description |
|------------|-------|-------|--------|----|--------|--------------------|
| | Score | Match | Length | | | |
| 1 | 1122 | 100.0 | 215 | 4 | 075303 | 075303 homo sapien |
| 2 | 947 | 84.4 | 352 | 6 | 095NC5 | 095NC5 hylobates s |
| 3 | 945 | 84.2 | 352 | 6 | 018772 | 018772 pan troglod |
| 4 | 940 | 83.8 | 352 | 6 | 018771 | 018771 pan troglod |
| 5 | 940 | 83.8 | 352 | 6 | 095NC0 | 095NC0 hylobates m |
| 6 | 939 | 83.7 | 352 | 6 | 09XS99 | 09XS99 gorilla gor |
| 7 | 938 | 83.6 | 352 | 6 | 09YT50 | 09YT50 pan troglod |
| 8 | 938 | 83.6 | 352 | 6 | 095NC8 | 095NC8 colobus pol |
| 9 | 938 | 83.6 | 352 | 6 | 095NC6 | 095NC6 trachypithe |
| 10 | 935 | 83.3 | 352 | 6 | 09MZ33 | 09MZ33 hylobates a |
| 11 | 935 | 83.3 | 352 | 6 | 09Y962 | 09Y962 pygathrix a |
| 12 | 934 | 83.2 | 352 | 6 | 09XT14 | 09XT14 colobus gue |
| 13 | 934 | 83.2 | 352 | 6 | 095NC7 | 095NC7 nasalis lar |
| 14 | 934 | 83.2 | 352 | 6 | 095NC3 | 095NC3 miopithecus |
| 15 | 933 | 83.2 | 352 | 6 | 018770 | 018770 pan troglod |
| 16 | 933 | 83.2 | 352 | 6 | 09YT49 | 09YT49 percocebus |

QY 121 LITDRLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LITDRLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HFPYKDSHGAPAAACHGHLGNPKNSASYSK 215
 DB 181 HFPYKDSHGAPAAACHGHLGNPKNSASYSK 215

RESULT 2

Q95NC5 PRELIMINARY; PRT; 352 AA.

AC Q95NC5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=9590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate
 RT phylogeny."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177884; AAK43367.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40508 MW; F4F64B3AD5AF658A CRC64;

Query Match 84.4%; Score 947; DB 6; Length 352;
 Best Local Similarity 98.4%; Pred. No. 1.9e-80;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAAQWDFGNTMCLLTGLYFGFSGIFFII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAAQWDFGNTMCLLTGLYFGFSGIFFII 120
 QY 121 LITDRLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LITDRLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 3

ID 018772 PRELIMINARY; PRT; 352 AA.

AC 018772;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CCR5 receptor (fragment).
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CHCCR5-142a;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
 RL EMBL; AF011541; AAB65741.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL2; 1.
 KW Receptor.
 FT NON_TER 352 352
 SQ SEQUENCE 352 AA; 40598 MW; 39688FA7004C952F CRC64;

Query Match 84.2%; Score 945; DB 6; Length 352;
 Best Local Similarity 98.4%; Pred. No. 2.9e-80;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAAQWDFGNTMCLLTGLYFGFSGIFFII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAAQWDFGNTMCLLTGLYFGFSGIFFII 120
 QY 121 LITDRLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LITDRLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 4

ID 018771 PRELIMINARY; PRT; 352 AA.

AC 018771;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CCR5 receptor (fragment).
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CHCCR5-141a;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
 RL EMBL; AF011539; AAB65739.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_FL2; 1.
 KW Receptor.
 FT NON_TER 352 352
 SQ SEQUENCE 352 AA; 40466 MW; 3EFFAC7ABAE1D4FB CRC64;

Query Match 83.8%; Score 940; DB 6; Length 352;
 Best Local Similarity 97.8%; Pred. No. 8.6e-80;
 Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60

```

QY 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 1 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 5
Q95NC0 PRELIMINARY; PRT; 352 AA.
AC Q95NC0;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Hylobates moloch (silvery gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=81572;
RN [1]
RP SEQUENCE FROM N.A.
RA "Sequence comparison of the CCR5 gene in primates and primate phylogeny."
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF177899; AAK43382.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40436 MW; 9623CA98340CF274 CRC64;

Query Match 83.8%; Score 940; DB 6; Length 352;
Best Local Similarity 97.8%; Pred. No. 8.6e-80;
Matches 180; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMLVILLINCKR 60
DB 1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 6
Q9XS99 PRELIMINARY; PRT; 352 AA.
AC Q9XS99;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GORILLACCR;
RX MEDLINE=99210133; PubMed=10195758;
RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T., Lal R.B.;
RT "Species-specific changes in the CCR5 gene from African and Asian nonhuman primates."
RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
DR EMBL; AF105291; AAD20560.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodpsn.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40529 MW; 1BB68C68FE2E7AD0 CRC64;

Query Match 83.78%; Score 939; DB 6; Length 352;
Best Local Similarity 97.3%; Pred. No. 1.1e-79;
Matches 179; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMLVILLINCKR 60
DB 1 MDYQVSSPTDIDYITSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
DB 121 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLGLYFIFGFSGIFII 120
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 7
Q9TV50 PRELIMINARY; PRT; 352 AA.
AC Q9TV50;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1410;
RX MEDLINE=99335215; PubMed=10408730;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with HIV carrier status in African nonhuman primates."
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1410;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.C., Barre-Sinoussi F., Fomsgaard A.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF035214; AAD44007.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodpsn.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; UNKNOWN_1.

```

DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07AA65 CRC64;
Query Match 83.6%; Score 938; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 1.3e-79;
Matches 178; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINYSSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
DB 1 MDYQSSPTDYDIDYTSSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFI 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFI 120
QY 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 8
Q95NC8 PRELIMINARY; PRT; 352 AA.
AC Q95NC8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Colobus polykomos.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Colobus.
OX NCBI_TaxID=9572;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177881; AAK43364.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40578 MW; 4366F149C3B4938F CRC64;
Query Match 83.6%; Score 938; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 1.3e-79;
Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINYSSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
DB 1 MDYQSSPTDYDIDYTSSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFI 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFI 120
QY 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 9
Q95NC6 PRELIMINARY; PRT; 352 AA.
AC Q95NC6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Trachypithecus johnii (hooded leaf monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Trachypithecus.
OX NCBI_TaxID=66063;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177883; AAK43366.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEPT_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEPT_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40462 MW; 52824E032259F7F CRC64;
Query Match 83.6%; Score 938; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 1.3e-79;
Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINYSSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
DB 1 MDYQSSPTDYDIDYTSSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFI 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFI 120
QY 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DB 121 LITIDRYLAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184
RESULT 10
Q95NC3 PRELIMINARY; PRT; 352 AA.
AC Q95NC3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CC chemokine receptor 5.
GN CCR5.
OS Hylobates agilis unko.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9583;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20317091; PubMed-10747879;
RA Mummidi S., Bamshad M., Ahuja S.S., Gonzalez E., Feuillet P.M.,
RA Begum K., Galvis M.C., Kosticki V., Valente A.J., Murthy K.K.,
RA Haro L., Dolan M.J., Allan J.S., Ahuja S.K.;
RT "Evolution of human and non-human primate CC chemokine receptor 5 gene
differential haplotype-specific transcriptional activity, and altered
transcription factor binding to polymorphic nucleotides in the

RT pathogenesis of HIV-1 and simian immunodeficiency virus.";

RL J. Biol. Chem. 275:18946-18961(2000).

DR EMBL; AF252551; AAF87981.1; -

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm.1; 1

DR PRINTS; PR00237; GPCRHHODPSN

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

SQ SEQUENCE 352 AA; 40293 MW; BAF8279ABAB5309 CRC64;

Query Match 83.3%; Score 935; DB 6; Length 352;

Best Local Similarity 96.7%; Pred. No. 2.5e-79;

Matches 178; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MDQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILLINCKR 60

DB 1 MDQVSSPTDIDYTSGPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILLINCKR 60

QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGIFII 120

DB 61 LKSMTDVYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGIFII 120

QY 121 LTTDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

DB 121 LTTDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

RESULT 11

O97962

ID O97962 PRELIMINARY; PRT; 352 AA.

AC O97962;

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE C-C chemokine receptor type 5.

GN CCR5.

OS Pygathrix avunculus (Tonkin snub-nosed monkey), and

OS Pygathrix roxellana (golden snub-nosed monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;

OC Pygathrix.

OC NCBI_TaxID=66062, 61622;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhang Y.-W., Zhang Y.-P.;

RT "Sequence evolution of chemokine receptor CCR5 gene in primates.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF075447; AAD19859.1; -

DR EMBL; AF075444; AAD19856.1; -

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCRHHODPSN

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

SQ SEQUENCE 352 AA; 40482 MW; 037CFA9E12E532F3 CRC64;

Query Match 83.3%; Score 935; DB 6; Length 352;

Best Local Similarity 95.7%; Pred. No. 2.5e-79;

Matches 176; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILLINCKR 60

DB 1 MDQVSSPTDIDYTSGPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILLINCKR 60

QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGIFII 120

DB 61 LKSMTDVYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGIFII 120

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 121 LTTDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

DB 121 LTTDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

RESULT 12

O9XT14

ID O9XT14 PRELIMINARY; PRT; 352 AA.

AC O9XT14;

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CC chemokine receptor 5.

GN CCR5.

OS Colobus guereza (Black-and-white colobus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;

OC Colobus.

OC NCBI_TaxID=33548;

RN [1]

RP SEQUENCE FROM N.A.

RA Spearman P.W., Mburu D.N., Graham B.S.;

RT "Differential utilization of CCR5 molecules from three East African

simian species by the HIV-1 envelope glycoprotein.";

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF141639; AAD32684.1; -

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCRHHODPSN

DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

SQ SEQUENCE 352 AA; 40550 MW; 9B078EF04D34DB36 CRC64;

Query Match 83.2%; Score 934; DB 6; Length 352;

Best Local Similarity 95.7%; Pred. No. 3.1e-79;

Matches 176; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDQVSSPIYDINVTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILLINCKR 60

DB 1 MDQVSSPTDIDYTSGPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILLINCKR 60

QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGIFII 120

DB 61 LKSMTDVYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCQLLTGLYFIFGFGIFII 120

QY 121 LTTDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

DB 121 LTTDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

QY 181 HFPY 184

DB 181 HFPY 184

```

OX NCBI_TaxID=43780;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RL phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177882; AAK43365.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40537 MW; 51F6F1486E35938E CRC64;

Query Match 83.2%; Score 934; DB 6; Length 352;
Best Local Similarity 95.7%; Pred. No. 3.1e-79;
Matches 176; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPTDIDYNTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNLVILLINCKR 60
DB 1 MDYQSSPTDIDYNTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFFII 120
QY 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 14
Q95NC3 PRELIMINARY; PRT; 352 AA.
AC Q95NC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Miopithecus.
OX NCBI_TaxID=36231;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RL phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177886; AAK43369.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40546 MW; 6464152F3E566AE5 CRC64;

Query Match 83.2%; Score 934; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 3.1e-79;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQSSPTDIDYNTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNLVILLINCKR 60
DB 1 MDYQSSPTDIDYNTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFFII 120

```

```

DB 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFFII 120
QY 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 15
O18770 PRELIMINARY; PRT; 352 AA.
ID O18770;
AC O18770;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CCR5 receptor (fragment).
GN CCR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MACCR5-140A;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT "HIV-1 subtypes, co-receptor usage, and CCR5 polymorphism.";
RL AIDS Res. Hum. Retroviruses 0:0-0(1997).
DR EMBL: AF011538; AAB65738.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsin.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 352 352
SQ SEQUENCE 352 AA; 40523 MW; 4513DB983A28ACH2 CRC64;

Query Match 83.2%; Score 933; DB 6; Length 352;
Best Local Similarity 96.2%; Pred. No. 3.8e-79;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQSSPTDIDYNTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNLVILLINCKR 60
DB 1 MDYQSSPTDIDYNTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFLLTPVFWAHYAAQWDFGNTMCLLTGLYFGFSGIFFII 120
QY 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LTTIDRYLAVHAFALKARTVTFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

Search completed: June 3, 2003, 19:23:53
Job time : 44.5732 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:18:46 ; Search time 12.2358 Seconds
(without alignments)
728.798 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122

Sequence: 1 MDYQVSSPIYDINTYSEPC.....AACGHLLGNKNSASVSK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 958 | 85.4 | 352 | 1 | CCR5_HUMAN |
| 2 | 952 | 84.8 | 352 | 1 | CCR5_PANTR |
| 3 | 947 | 84.4 | 352 | 1 | CCR5_PORPY |
| 4 | 942 | 84.0 | 352 | 1 | CCR5_GORGO |
| 5 | 938 | 83.6 | 352 | 1 | CCR5_PYGBI |
| 6 | 938 | 83.6 | 352 | 1 | CCR5_PYGNE |
| 7 | 938 | 83.6 | 352 | 1 | CCR5_TRAFR |
| 8 | 936 | 83.6 | 352 | 1 | CCR5_TRAPH |
| 9 | 936 | 83.4 | 352 | 1 | CCR5_HYLLI |
| 10 | 933 | 83.2 | 352 | 1 | CCR5_MACMU |
| 11 | 933 | 83.2 | 352 | 1 | CCR5_PAPHA |
| 12 | 928 | 82.7 | 352 | 1 | CCR5_CERTO |
| 13 | 916 | 81.6 | 352 | 1 | CCR5_CERAE |
| 14 | 786 | 70.1 | 354 | 1 | CCR5_MOUSE |
| 15 | 775 | 69.1 | 354 | 1 | CCR5_RAT |
| 16 | 715.5 | 63.8 | 373 | 1 | CCR2_MOUSE |
| 17 | 714.5 | 63.7 | 373 | 1 | CCR2_RAT |
| 18 | 694 | 61.9 | 374 | 1 | CCR2_HUMAN |
| 19 | 688 | 61.3 | 360 | 1 | CCR2_MACMU |
| 20 | 598 | 53.3 | 355 | 1 | CCR1_HUMAN |
| 21 | 570.5 | 50.8 | 355 | 1 | CCR1_MOUSE |
| 22 | 562 | 50.1 | 355 | 1 | CCR1_MACMU |
| 23 | 547.5 | 48.8 | 360 | 1 | CCR4_HUMAN |
| 24 | 546.5 | 48.7 | 360 | 1 | CCR4_MOUSE |
| 25 | 535.5 | 47.7 | 359 | 1 | CCR3_MOUSE |
| 26 | 535.5 | 47.7 | 359 | 1 | CCR3_RAT |
| 27 | 516.5 | 46.0 | 358 | 1 | CCR3_CAVPO |
| 28 | 506.5 | 45.1 | 355 | 1 | CCR3_HUMAN |
| 29 | 499.5 | 44.5 | 355 | 1 | CCR3_MACMU |
| 30 | 489.5 | 43.6 | 355 | 1 | CCR3_CERAE |
| 31 | 489.5 | 43.6 | 356 | 1 | CCR3_MOUSE |
| 32 | 431 | 38.4 | 353 | 1 | CCR8_MOUSE |
| 33 | 427.5 | 38.1 | 355 | 1 | CCR8_HUMAN |

RESULT 1
CCR5_HUMAN STANDARD; PRT: 352 AA.
AC P51681; O14692; O14693; O14695; O14696; O14697; O14698; O14699;
AC O14700; O14701; O14702; O14703; O14704; O14705; O14706; O14707;
AC O14708; O15538; Q9UPA4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5)
DE (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen).
GN CCR5 OR CCR5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96241590; PubMed-8639485;
RA Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.;
RT "Molecular cloning and functional expression of a new human
CC-chemokine receptor gene."
RL Biochemistry 35:3362-3367(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-96291862; PubMed-8663314;
RA Raport C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.;
RT "Molecular cloning and functional characterization of a novel human
CC chemokine receptor (CCR5) for RANTES, MIP-beta, and MIP-1alpha."
RL J. Biol. Chem. 271:17161-17166(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96295970; PubMed-8699119;
RA Combadere C., Ahuja S.K., Tiffany H.L., Murphy P.M.;
RT "Cloning and functional expression of CC CCR5, a human monocyte CC
chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and
RANTES."
RL J. Leukoc. Biol. 60:147-152(1996).
RN [4]
RP SEQUENCE FROM N.A.
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,
Nhan M., Farnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,
Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J.,
Muzny D., Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M.,
Porcel B.M., Dragan Y., Giacalone J., Pae A., Powell E.,
Solinsky K.A., Desilva U., Diaz-Perez S., Zhou X., Yu Y.,
Watanabe M., Doggett N., Garcia D., Sagripanti J.L.;
RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-98001387; PubMed-9343222;
RA Kuhnmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
implicate specific amino acids in infections by simian and human
immunodeficiency viruses."
RL J. Virol. 71:8642-8656(1997).

097665 macaca mula
P35411 rattus norv
Q920d9 mus musculu
Q9wt7 mus musculu
P49238 homo sapien
P51686 homo sapien
P32248 homo sapien
P47774 mus musculu
O62747 cercobebus
P30991 homo sapien
P79394 macaca mula
O00590 homo sapien

ALIGNMENTS

34 424 37.8 356 1 CCR8_MACMU
35 386 34.4 354 1 C3X1_RAT
36 383 34.1 354 1 C3X1_MOUSE
37 370 33.0 369 1 CCR9_MOUSE
38 365 32.5 355 1 C3X1_HUMAN
39 363 32.4 357 1 CCR9_HUMAN
40 345.5 30.8 378 1 CCR7_HUMAN
41 342.5 30.5 378 1 CCR7_MOUSE
42 334.5 29.8 352 1 CCR4_CERTO
43 331.5 29.5 352 1 CCR4_HUMAN
44 331.5 29.5 352 1 CCR4_MACMU
45 330.5 29.5 384 1 CCR6_HUMAN

[6]
 RP SEQUENCE FROM N.A., AND POLYMORPHISMS.
 RX MEDLINE-98022612; PubMed-9359654;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98049523; PubMed-9388201;
 RA Mummidi S., Ahuja S.S., McDaniel B.L., Ahuja S.K.;
 RT "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts
 with 5'-end heterogeneity, dual promoter usage, and evidence for
 polymorphisms within the regulatory regions and noncoding exons.";
 RN J. Biol. Chem. 272:30662-30671(1997).
 [8]
 RP SEQUENCE FROM N.A., AND VARIANT ARG-178.
 RA Magierowska M., Barre-Sinoussi F., Issafras H., Theodorou I.,
 RA Debre P.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 RX MEDLINE-96260017; PubMed-8649511;
 RA Deng H., Liu R., Ellmeier W., Choe S., Unutmaz D., Burkhardt M.,
 RA di Marzio P., Marmon S., Sutton R.E., Hill C.M., Davis C.B.,
 RA Peiper S.C., Schall T.J., Littman D.R., Landau N.R.;
 RT "Identification of a major co-receptor for primary isolates of
 RT HIV-1.";
 RL Nature 381:661-666(1996).
 [10]
 RP CHARACTERIZATION OF ITS HIV-1 RECEPTOR FUNCTION.
 RX MEDLINE-96260018; PubMed-8649512;
 RA Dragic T., Litwin V., Allaway G.P., Martin S.R., Huang Y.,
 RA Nagashima K.A., Cayanan C., Maddon P.J., Koup R.A., Moore J.P.,
 RA Paxton W.A.;
 RT "HIV-1 entry into CD4+ cells is mediated by the chemokine receptor
 RT CC-CKR-5.";
 RL Nature 381:667-673(1996).
 [11]
 RP SULFATION.
 RX MEDLINE-99189752; PubMed-10089882;
 RA Farzan M., Mirzabekov T., Kolchinsky P., Wyatt R., Cayabyab M.,
 RA Gerard N.P., Gerard C., Sodroski J., Choe H.;
 RT "Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1
 RT entry.";
 RL Cell 96:667-676(1999).
 CC -!- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION. ACTS AS CO-RECEPTOR WITH CD4 FOR PRIMARY NON-
 CC SYNCTIOM-INDUCING STRAINS (NSI) (MACROPHAGE-TROPIC) OF HIV-1
 CC VIRUS. IT PROMOTES ENV-MEDIATED FUSION OF THE VIRUS.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: FOUND IN PROMYELOCYTIC CELLS.
 CC -!- PTM: SULFATION CONTRIBUTES TO THE EFFICIENCY OF HIV-1 ENTRY.
 CC -!- PTM: MODIFIED BY O-LINKED GLYCOSYLATION, BUT NOT BY N-LINKED
 CC GLYCOSYLATION.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC -----
 DR EMBL; X91492; CAA62796.1; -
 DR EMBL; U54994; AAC50598.1; -
 DR EMBL; U57840; AAB17071.1; -
 DR EMBL; U95626; AAB57793.1; -
 DR EMBL; U83326; AAC51797.1; -

DR EMBL; AF011500; AAB65700.1; -
 DR EMBL; AF011501; AAB65701.1; -
 DR EMBL; AF011502; AAB65702.1; -
 DR EMBL; AF011503; AAB65703.1; -
 DR EMBL; AF011505; AAB65705.1; -
 DR EMBL; AF011506; AAB65706.1; -
 DR EMBL; AF011507; AAB65707.1; -
 DR EMBL; AF011508; AAB65708.1; -
 DR EMBL; AF011509; AAB65709.1; -
 DR EMBL; AF011510; AAB65710.1; -
 DR EMBL; AF011511; AAB65711.1; -
 DR EMBL; AF011512; AAB65712.1; -
 DR EMBL; AF011513; AAB65713.1; -
 DR EMBL; AF011514; AAB65714.1; -
 DR EMBL; AF011515; AAB65715.1; -
 DR EMBL; AF011516; AAB65716.1; -
 DR EMBL; AF011517; AAB65717.1; -
 DR EMBL; AF011518; AAB65718.1; -
 DR EMBL; AF011519; AAB65719.1; -
 DR EMBL; AF011520; AAB65720.1; -
 DR EMBL; AF011521; AAB65721.1; -
 DR EMBL; AF011522; AAB65722.1; -
 DR EMBL; AF011523; AAB65723.1; -
 DR EMBL; AF011524; AAB65724.1; -
 DR EMBL; AF011525; AAB65725.1; -
 DR EMBL; AF011526; AAB65726.1; -
 DR EMBL; AF011527; AAB65727.1; -
 DR EMBL; AF011528; AAB65728.1; -
 DR EMBL; AF011529; AAB65729.1; -
 DR EMBL; AF011530; AAB65730.1; -
 DR EMBL; AF011531; AAB65731.1; -
 DR EMBL; AF011532; AAB65732.1; -
 DR EMBL; AF011533; AAB65733.1; -
 DR EMBL; AF011534; AAB65734.1; -
 DR EMBL; AF011535; AAB65735.1; -
 DR EMBL; AF011536; AAB65736.1; -
 DR EMBL; AF011537; AAB65737.1; -
 DR EMBL; AF031237; AAB94735.1; -
 DR EMBL; AF052539; AAD18131.1; -
 DR Genew; HGNC:1606; CCR5.
 DR MIM; 601373; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PR00237; G-PROTEIN RECP_FL_1; 1.
 DR PROSITE; PS50262; G-PROTEIN RECP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
 KW Polymorphism.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION.
 FT MOD_RES 10 10 SULFATION.
 FT MOD_RES 14 14 SULFATION.
 FT MOD_RES 15 15 SULFATION.
 FT VARIANT 10 10 Y -> D (IN INCCR5-71A).
 FT VARIANT 29 29 A -> S (IN DESNP:1800939).
 FT VARIANT 31 31 R -> H (IN INCCR5-72A).

```
Query Match      85.4%; Score 958; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.3e-53;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSPYIDNYITSEPCQKINVKQIAARLLPPLSLVFIKFGVGNMVLILLINCKR 60
    |||||
Db 1 MDYQVSPYIDNYITSEPCQKINVKQIAARLLPPLSLVFIKFGVGNMVLILLINCKR 60
    |||||

QY 61 LKSMTDIYLNLAIISDLFFLLTPFVAHYAAQWDFGNTMCQLLTGLYIFGFGSIFII 120
    |||||
Db 61 LKSMTDIYLNLAIISDLFFLLTPFVAHYAAQWDFGNTMCQLLTGLYIFGFGSIFII 120
    |||||

QY 121 LITIDRYLAVVHAFALKARTVFGVVTSTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
    |||||
Db 121 LITIDRYLAVVHAFALKARTVFGVVTSTVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
    |||||

QY 181 HFPY 184
    ||||
Db 181 HFPY 184

RESULT 2
CRK5_PANTR STANDARD; PRT; 352 AA.
ID CRK5_PANTR AC P56440; 002778;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CRK-5) (CCR-5) (CCR5).
GN CCR5 OR CMKBR5.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharon M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
RA Pelper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT Differential utilization of CCR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains.;
RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
RN [2]
RN SEQUENCE FROM N.A.
RA Zimmerman P.A., Buckler-White A., Alkhatib G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=98022612; PubMed=9359654;
RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
RA Ho D.D.;
RT HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.;
RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=97426118; PubMed=9282822;
RA Zacharova V., Zachar V., Goustin A.S.;
RT Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural
RT HIV type 1 host.;
RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).
RN [5]
RN SEQUENCE FROM N.A.
RX MEDLINE=98090115; PubMed=9430250;
RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;
RT Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.;
RN AIDS Res. Hum. Retroviruses 13:1583-1587(1997).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
```


30-MAY-2000 (Rel. 39, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
 CCR5 OR CMKR5.
 Pongo pygmaeus (Orangutan).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 NCBI_TaxID=9600;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=99416438; PubMed=10486970;
 Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 Mol. Biol. Evol. 16:1145-1154(1999).
 -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 DIFFERENTIATION.
 -1- SUBCELLULAR LOCATION: Integral membrane protein.
 -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 EMBL; AF075446; AAD19858.1; -
 InterPro; IPR000276; GPCR_Rhodpsn.
 Pfam; PF00001; 7tm1.1;
 PRINTS; PR00237; GPCRHHODPSN.
 PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 FT MOD_RES 352 AA; 40527 MW; F4E2F47135AF658A CRC64;
 SQ SEQUENCE 352 AA; 40527 MW; F4E2F47135AF658A CRC64;
 Query Match 84.4%; Score 947; DB 1; Length 352;
 Best Local Similarity 98.4%; Pred. No. 1.le-52;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDVQSSPIYDINVTSEPCQKINVKQIAARLLPPLSLVIFGFGVGNMVLILLNCKR 60
 Db 1 MDVQSSPIYDINVTSEPCQKINVKQIAARLLPPLSLVIFGFGVGNMVLILLNCKR 60
 QY 61 LKSWTDIYLNLAIISDLFELLTPVFWAHYAAQDFGNTMCLLGLYFGFGFIIFII 120
 Db 61 LKSWTDIYLNLAIISDLFELLTPVFWAHYAAQDFGNTMCLLGLYFGFGFIIFII 120
 QY 121 LLTIDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLITFRSQEGLHYTCSS 180
 Db 121 LLTIDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLITFRSQEGLHYTCSS 180

Db 121 LLTIDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLITFRSQEGLHYTCSS 180
 QY 181 HFPY 184
 Db 181 HFPY 184
 RESULT 4
 CCR5_GORGO
 ID CCR5_GORGO STANDARD; PRT; 352 AA.
 AC P56439;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5).
 GN CCR5 OR CMKR5.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OC NCBI_TaxID=9595;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 simian immunodeficiency virus strains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 EMBL; AF005659; AAB62553.1; -
 InterPro; IPR000276; GPCR_Rhodpsn.
 Pfam; PF00001; 7tm1.1;
 PRINTS; PR00237; GPCRHHODPSN.
 PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 FT MOD_RES 352 AA; 40515 MW; D0E6FCB9FE5EAC84 CRC64;
 SQ SEQUENCE 352 AA; 40515 MW; D0E6FCB9FE5EAC84 CRC64;

```
Query Match      84.0%; Score 942; DB 1; Length 352;
Best Local Similarity 97.8%; Pred. No. 2.3e-52;
Matches 180; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINTTSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MDYQSSPTDIDYTTSEPCQKTNVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 LTTIDRYLAVHAFALKARTVGVVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 LTTIDRYLAVHAFALKARTVGVVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 181 HPFY 184
    ||||
DB 181 HPFY 184
    ||||

RESULT 5
CKR5_PYGBI
ID CKR5_PYGBI STANDARD; PRT; 352 AA.
AC 097860;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
GN CKR5 OR CMKR5.
OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.
OX NCBI_TaxID=61621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99416438; PubMed=10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CKR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL; AF075445; AAD19857.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm1.1; 1.
CC PRINTS; PR00237; GPCRHOODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 59 68
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT TRANSMEM 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 4366F142730F938F CRC64;

Query Match      83.6%; Score 938; DB 1; Length 352;
Best Local Similarity 96.2%; Pred. No. 4e-52;
Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINTTSEPCQKINVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MDYQSSPTDIDYTTSEPCQKTNVKQIAARLLPPLYSIVFIFGVGNMVLILLINCKR 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 LKSMTDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCLLTGLYFIFGFSGIFPII 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 LTTIDRYLAVHAFALKARTVGVVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 LTTIDRYLAVHAFALKARTVGVVTSVITWVAVFASLPGIIFTRSQEGLHYTCSS 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 181 HPFY 184
    ||||
DB 181 HPFY 184
    ||||

RESULT 6
CKR5_PYGNE
ID CKR5_PYGNE STANDARD; PRT; 352 AA.
AC 097882;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN CKR5 OR CMKR5.
OS Pygathrix nemaeus (Dove langur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Pygathrix.
OX NCBI_TaxID=54133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99416438; PubMed=10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CKR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
CC EMBL; AF075448; AAD19860.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
```

PFam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G-PROTEIN RECF_F1_1; 1.
 DR PROSITE: PS00237; G-PROTEIN RECF_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40532 MW; FE4F9D98D3B3E861 CRC64;
 Query Match 83.6%; Score 938; DB 1; Length 352;
 Best Local Similarity 96.2%; Pred. No. 4e-52;
 Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYVSSPTDYDNTYTPCQKINVKQIAARLLPPLYSIVTFGFGNMLVILLINCR 60
 DB 1 MDYVSSPTDYDNTYTPCQKINVKQIAARLLPPLYSIVTFGFGNMLVILLINCR 60
 QY 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFII 120
 DB 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFII 120
 QY 121 LLTIDRYLVAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LLTIDRYLVAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184
 RESULT 7
 CKR5_TRAFR STANDARD; PRT; 352 AA.
 AC 097878;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CMKR5.
 OS Trachypithecus francoisi (Francois' langur) (Indochinese langur).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
 OC Trachypithecus.
 OX NCBI_TaxID=54180;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.

-1- SUBCELLULAR LOCATION: Integral membrane protein.
 -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF075442; AAD19854.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G-PROTEIN RECF_F1_1; 1.
 DR PROSITE: PS00237; G-PROTEIN RECF_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40496 MW; 4366F148C25938F CRC64;
 Query Match 83.6%; Score 938; DB 1; Length 352;
 Best Local Similarity 96.2%; Pred. No. 4e-52;
 Matches 177; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDYVSSPTDYDNTYTPCQKINVKQIAARLLPPLYSIVTFGFGNMLVILLINCR 60
 DB 1 MDYVSSPTDYDNTYTPCQKINVKQIAARLLPPLYSIVTFGFGNMLVILLINCR 60
 QY 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFII 120
 DB 61 LKSMTDIYLNLAIISDLFFLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFGSIFII 120
 QY 121 LLTIDRYLVAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LLTIDRYLVAVHAFALKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 DB 181 HPFY 184
 RESULT 8
 CKR5_TRAPH STANDARD; PRT; 352 AA.
 AC 097879;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CMKR5.
 OS Trachypithecus phayrei (Phayre's leaf monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;
 OC Trachypithecus.


```

Db 61 LKSMTDIYLNLAIASDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFGFGIFPII 120
QY 121 LITDRYLAVHVAFAVKARTVGVVTSVITWVAVFASLPGLIIFTSQKGLHYTCSS 180
Db 121 LITDRYLAVHVAFAVKARTVGVVTSVITWVAVFASLPGLIIFTSQKGLHYTCSS 180
QY 181 HPFY 184
Db 181 HPFY 184

RESULT 10
CKR5_MACMU STANDARD; PRT; 352 AA.
ID CKR5_MACMU
AC P79436; 002746;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN CKR5 OR CKMR5.
OS Macaca mulatta (Rhesus macaque).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca nemestrina (Pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544, 9541, 9545;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=97184592; PubMed=9032394;
RA Marcon L., Choe H., Martin K.A., Farzan M., Ponath P.D., Wu L.,
RA Newman W., Gerard N., Gerard C., Sodroski J.;
RT Utilization of C-C chemokine receptor 5 by the envelope
RT glycoproteins of a pathogenic simian immunodeficiency virus,
RT SIVmac239.
RT J. Virol. 71:2522-2527 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=97213934; PubMed=9060623;
RA Chen Z., Zhou P., Ho D.D., Landau N.R., Marx P.A.;
RT "Genetically divergent strains of simian immunodeficiency virus use
RT CKR5 as a coreceptor for entry."
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RX MEDLINE=21354176; PubMed=11461684;
RA Margulies B.J., Hauer D.A., Clements J.E.;
RT "Identification and comparison of eleven rhesus macaque chemokine
RT receptors."
RN [4]
RP AIDS Res. Hum. Retroviruses 17:981-986 (2001).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta, M.fascicularis, and M.nemestrina;
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharon M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
RT "Differential utilization of CKR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains."
RN [6]
RP Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010 (1997).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U77672; AAC511109.1; -
DR EMBL; U73739; AAC511158.1; -
DR EMBL; U96762; AAC34132.1; -
DR EMBL; AF005660; AB62554.1; -
DR EMBL; AF005661; AB62555.1; -
DR EMBL; AF005662; AB62556.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1.1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 241 241 M -> I (IN REF. 3).
FT CONFLICT 292 292 I -> M (IN REF. 3).
SQ SEQUENCE 352 AA; 40507 MW; 58B96C85909FACB2 CRC64;

Query Match 83.2%; Score 933; DB 1; Length 352;
Best Local Similarity 96.2%; Pred. No. 8.3e-52;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDYQVSPYDINYYTSBPCKINVKQIAARLLPPLYSLVFIFGVGNMVLILINCKR 60
Db 1 MDYQVSPYDIDYTSBPCKINVKQIAARLLPPLYSLVFIFGVGNMVLILINCKR 60
QY 61 LKSMTDIYLNLAIASDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFGFGIFPII 120
Db 61 LKSMTDIYLNLAIASDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFGFGIFPII 120
QY 121 LITDRYLAVHVAFAVKARTVGVVTSVITWVAVFASLPGLIIFTSQKGLHYTCSS 180
Db 121 LITDRYLAVHVAFAVKARTVGVVTSVITWVAVFASLPGLIIFTSQKGLHYTCSS 180
QY 181 HPFY 184
Db 181 HPFY 184

RESULT 11
CKR5_PAPHA STANDARD; PRT; 352 AA.
ID CKR5_PAPHA
AC P56441;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).

```

GN CCR5 OR CMKR5.
OS Papio hamadryas (Hamadryas baboon), and
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9557, 9555;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-P.hamadryas;
RX MEDLINE=97268687; PubMed=9108095;
RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
RA Pelier S.C., Parmentier M., Broder C.C., Doms R.W.;
RT "Differential utilization of CCR5 by macrophage and T cell tropic
RT simian immunodeficiency virus strains."
RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-P.hamadryas;
RX MEDLINE=99210133; PubMed=10195758;
RA Sakuna N.K., Wang B., Novembre F.J., Bolton W., Smit T.K., Lal R.B.;
RT "Species-specific changes in the CCR5 gene from African and Asian
RT nonhuman primates."
RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-P.anubis;
RA Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF005658; AAB62552.1; -
CC EMBL; AF105287; AAD20556.1; -
CC EMBL; AF105288; AAD20557.1; -
CC EMBL; AF105289; AAD20558.1; -
CC EMBL; AF105290; AAD20559.1; -
CC EMBL; AF023452; AAC63830.1; -
CC Pfam; PF00001; 7tm_1; 1.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30
FT 1 EXTRACELLULAR (POTENTIAL).
FT 31 58
FT 1 CYTOPLASMIC (POTENTIAL).
FT 59 68
FT 1 TRANSMEM
FT 69 89
FT 2 (POTENTIAL).
FT 90 102
FT 1 EXTRACELLULAR (POTENTIAL).
FT 103 124
FT 3 (POTENTIAL).
FT 125 141
FT 1 CYTOPLASMIC (POTENTIAL).
FT 142 166
FT 4 (POTENTIAL).
FT 167 198
FT 1 EXTRACELLULAR (POTENTIAL).
FT 199 218
FT 5 (POTENTIAL).
FT 219 235
FT 6 CYTOPLASMIC (POTENTIAL).
FT 236 260
FT 6 (POTENTIAL).
FT 261 277
FT 1 EXTRACELLULAR (POTENTIAL).
FT 278 301
FT 7 (POTENTIAL).
FT 302 352
FT 1 CYTOPLASMIC (POTENTIAL).

FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
FT CARBOHYD 268 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 352 AA; 40489 MW; 5E1504A9BAJFEB2 CRC64;
Query Match 83.2%; Score 933; DB 1; Length 352;
Best Local Similarity 96.2%; Pred. No. 8.3e-52;
Matches 177; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDYVSSPIYDINYYSEPCQKINVKQIAARLLPPLYSLVFVFGVGNMVLILLINCKR 60
Db 1 MDYVSSPIYDINYYSEPCQKINVKQIAARLLPPLYSLVFVFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAAWDFGNTMCOLLTGLYFGFFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAAWDFGNTMCOLLTGLYFGFFSGIFII 120
QY 121 LLTIDRYLAHVAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTSOKEGLHYTCSS 180
Db 121 LLTIDRYLAHVAVFALKARTVTFGVVTSVITWVAVFASLPGLIFTSOKEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184
RESULT 12
ID_CKR5_CERTO STANDARD; PRT: 352 AA.
AC 062743; 062744; 062745; 062746;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-CR5) (CCR-5) (CCR5).
GN CCR5 OR CMKR5.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9531;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate 079, 085, 087, and 089;
RX MEDLINE=98321155; PubMed=9656999;
RA Chen Z., Gettie A., Ho D.D., Marx P.A.;
RT "Primary SIVsm isolates use the CCR5 coreceptor from sooty mangabeys
RT naturally infected in west Africa: a comparison of coreceptor usage
RT of primary SIVsm, HIV-2, and SIVmac."
RL Virology 246:113-124(1998).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF051902; AAC39830.1; -
CC EMBL; AF051903; AAC39831.1; -
CC EMBL; AF051904; AAC39832.1; -
CC EMBL; AF051905; AAC39833.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.

```

DR PFAM: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_HODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
FT DOMAIN 1 30
FT TRANSMEM 31 58
FT DOMAIN 59 68
FT TRANSMEM 69 89
FT DOMAIN 90 102
FT TRANSMEM 103 124
FT DOMAIN 125 141
FT TRANSMEM 142 166
FT DOMAIN 167 198
FT TRANSMEM 199 218
FT DOMAIN 219 235
FT TRANSMEM 236 260
FT DOMAIN 261 277
FT TRANSMEM 278 301
FT DOMAIN 302 325
FT TRANSMEM 326 352
FT DISULFID 101 178
FT MOD_RES 3 3
FT MOD_RES 10 10
FT MOD_RES 14 14
FT MOD_RES 15 15
FT VARIANT 2 2
FT VARIANT 3 3
FT VARIANT 25 25
FT VARIANT 100 100
FT VARIANT 107 107
FT VARIANT 134 134
FT VARIANT 146 146
FT VARIANT 340 340
FT SEQUENCE 352 AA; 40489 MW; 20A196E2D47E49CA CRC64;

Query Match 82.7%; Score 928; DB 1; Length 352;
Best Local Similarity 95.1%; Pred. No. 1.7e-51;
Matches 176; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
DB 1 MDYQVSSPTDIDYITSEPCQKINVKQIARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
QY 121 LLTIDRYLAVHVFALKARTVFGVTSVITWWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LLTIDRYLAVHVFALKARTVFGVTSVITWWVAVFASLPGLIIFTRSQEGLHYTCSP 180
QY 181 HEPY 184
DB 181 HEPY 184

RESULT 13
CKR5_CERAE
ID CKR5_CERAE STANDARD; PRT; 352 AA.
AC P56493;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
GN CCR5 OR CMKR5.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;

RX MEDLINE=98001387; PubMed=9343222;
RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Murayama Y., Matsunaga S., Inoue-Murayama M.;
RT "cDNA sequence of African green monkey CCR-5 chemokine receptor
RT gene.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC 1- SUBCELLULAR LOCATION: Integral membrane protein.
CC 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U83324; AAC51795.1; -.
CC EMBL: U83325; AAC51796.1; -.
CC EMBL: AB015944; BAA31328.1; -.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm.1; 1.
CC PRINTS: PR00237; GPCR_HODOPSIN.
CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
CC Polymorphism.
KW FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 31 58 1 (POTENTIAL).
KW DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
KW TRANSMEM 69 89 2 (POTENTIAL).
KW DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 103 124 3 (POTENTIAL).
KW DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
KW TRANSMEM 142 166 4 (POTENTIAL).
KW DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 199 218 5 (POTENTIAL).
KW DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
KW TRANSMEM 236 260 6 (POTENTIAL).
KW DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 278 301 7 (POTENTIAL).
KW DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
KW DISULFID 101 178 BY SIMILARITY.
KW MOD_RES 3 3 SULFATION (BY SIMILARITY).
KW MOD_RES 10 10 SULFATION (BY SIMILARITY).
KW MOD_RES 14 14 SULFATION (BY SIMILARITY).
KW VARIANT 2 2 N -> Y.
KW VARIANT 3 2 F -> L.
KW SEQUENCE 352 AA; 40561 MW; 7F52E690C72EC29A CRC64;

Query Match 81.6%; Score 916; DB 1; Length 352;
Best Local Similarity 95.1%; Pred. No. 9.4e-51;
Matches 175; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
DB 1 MDYQVSSPTDIDYITSEPCQKINVKQIARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLALSIDFLFTVPFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFII 120

```

QY 121 LFTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGIIFRSQREGLHTCSS 180
 DB 121 LFTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGIIFRSQREGLHTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 14
 CKR5_MOUSE STANDARD; PRT; 354 AA.
 AC P51682; Q61867; P97405; O35313; P97308; O35891;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (MIP-1
 alpha receptor).
 DE alpha receptor.
 GN CNR5 OR CNR5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SVJ; TISSUE=Spleen;
 RX MEDLINE=96205938; PubMed=6331787;
 RA Boring L., Gosling J., Monteclaro F.S., Lusis A.J., Tsou C.-L.,
 RA Chato I.F.;
 RT "Molecular cloning and functional expression of murine JE (monocyte
 chemoattractant protein 1) and murine macrophage inflammatory protein
 1alpha receptors: evidence for two closely linked C-C chemokine
 receptors on chromosome 9.";
 RT J. Biol. Chem. 271:7551-7558(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6 X CHA; TISSUE=Thymus;
 RX MEDLINE=96278910; PubMed=8662890;
 RA Meyer A., Coyne A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
 RT "Cloning and characterization of a novel murine macrophage
 inflammatory protein-1 alpha receptor.";
 RT J. Biol. Chem. 271:14445-14451(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/Ola;
 RA Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6, and NIH Swiss; TISSUE=Liver, Kidney, and Spleen;
 RX MEDLINE=98001387; PubMed=9343222;
 RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
 implicate specific amino acids in infections by simian and human
 immunodeficiency viruses.";
 RT J. Virol. 71:8642-8656(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129;
 RX MEDLINE=97404635; PubMed=9261347;
 RA Doranz B.J., Lu Z.H., Rucker J., Zhang T.Y., Sharron M., Cen Y.H.,
 RA Wang Z.X., Guo H.H., Du J.G., Accavitti M.A., Doms R.W., Peiper S.C.;
 RT "Two distinct CCR5 domains can mediate coreceptor usage by human
 immunodeficiency virus type 1.";
 RT J. Virol. 71:6305-6314(1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Guo B., Kuuo K., Harada A., Matsushima K.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: DETECTED IN MONOCYTE/MACROPHAGE CELL LINES,
 CC BUT NOT IN NONHEMATOPOIETIC CELL LINES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U47036; AAC52454.1; -;
 CC EMBL: X94151; CRA63867.1; -;
 CC EMBL: U86565; AAB37273.1; -;
 CC EMBL: U83327; AAC53386.1; -;
 CC EMBL: AF022990; AAC53389.1; -;
 CC EMBL: AF019772; AAB71183.1; -;
 CC EMBL: D83648; BAA12024.1; -;
 CC MCD; MGI:107182; Cnkr5.
 CC InterPro: IPR000276; GPCR_Rhodopsn.
 CC Pfam: PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHHODPSN.
 CC PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
 CC PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
 KW DOMAIN 1 32
 FT TRANSSEM 33 60
 FT DOMAIN 61 70
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 71 91
 FT DOMAIN 92 104
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 105 126
 FT DOMAIN 127 143
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 144 168
 FT DOMAIN 169 200
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 201 220
 FT DOMAIN 221 237
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 238 262
 FT DOMAIN 263 279
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 280 303
 FT DOMAIN 304 354
 FT CYTOPLASMIC (POTENTIAL).
 FT DISULFID 103 180
 FT CARBOHYD 270 270
 FT VARIANT 11 11 I -> S.
 FT VARIANT 62 62 K -> R.
 FT VARIANT 66 66 V -> M.
 FT VARIANT 97 97 I -> V.
 FT VARIANT 109 109 V -> L.
 FT VARIANT 156 156 V -> A.
 FT VARIANT 160 160 F -> S.
 FT VARIANT 185 185 P -> L.
 FT VARIANT 213 213 I -> V.
 FT VARIANT 318 318 I -> M.
 FT VARIANT 337 337 V -> A.
 FT CONFLICT 3 3 F -> L (IN REF. 2).
 FT CONFLICT 80 80 L -> F (IN REF. 2).
 FT CONFLICT 145 145 N -> I (IN REF. 5).
 FT CONFLICT 190 190 H -> Y (IN REF. 3).
 FT CONFLICT 208 208 P -> S (IN REF. 1).
 FT CONFLICT 208 208 P -> S (IN REF. 1).
 SQ SEQUENCE 354 AA; B4A6B942E88F9CF0 CRC64;
 Query Match 70.1%; Score 786; DB 1; Length 354;
 Best Local Similarity 80.1%; Pred. No. 1.le-42;
 Matches 149; Conservative 17; Mismatches 18; Indels 2; Gaps 1;
 QY 1 MDYVSSP--IYDINVTSEPCOKINVKQIAARLLPPLSLVFIKFGVGNMVLILLINC 58
 Db 1 MDEGVSPTIYDIDYGNAPCOKINVKQIAARLLPPLSLVFIKFGVGNMVLILLISC 60
 QY 59 KRLKSTDIYLLNLAIASDLFLTLTPFWAHYAQAQWDFGNMTCOLLTGLYFIFGSGIFP 118
 Db 61 KRLKSTDIYLLNLAIASDLFLTLTPFWAHYAQAQWDFGNMTCOLLTGLYFIFGSGIFP 120


```

QY 119 ILLTIDRYLAVHAFALKARTVTRGVTVITWVAVFASLPGLITRSQKGLHYTC 178
Db 121 ILLTIDRYLAVHAFALKARTVTRGVTVITWVAVFASLPGLITRSQKGLHYTC 180
QY 179 SSHPY 184
Db 181 SPHPH 186

RESULT 15
CKR5_RAT
ID CKR5_RAT STANDARD; PRT; 354 AA.
AC O08556;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (MIP-1
DE alpha receptor).
DE CCR5 OR CCR5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=98334064; PubMed=9670989;
RA Spleiss O., Gourmala N., Boddeke H.W.G.M., Sauter A., Fiebich B.L.,
RA Berger M., Gebicke-Haerter P.J.;
RT "Cloning of rat HIV-1-chemokine coreceptor CCR5 from microglia and
RT upregulation of its mRNA in ischemic and endotoxinemic rat brain.";
RL J. Neurosci. Res. 53:16-28(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=98318173; PubMed=9655467;
RA Jiang Y., Salafranca M.N., Adhikari S., Xia Y., Feng L., Sonntag M.K.,
RA Defebvre C.M., Pennell N.A., Streit W.J., Harrison J.K.;
RT "Chemokine receptor expression in cultured glia and rat experimental
RT allergic encephalomyelitis.";
RL J. Neuroimmunol. 86:1-12(1998).
CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Y12009; CA97237.1; -
CC DR EMBL; U77350; AAC03243.1; -
CC DR InterPro; IPR000276; GPCR_Rhodpsn.
CC DR Pfam; PF00001; 7tm_1; 1.
CC DR PRINTS; PR00237; GPCRHHODPSN.
CC DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 33 60 1 (POTENTIAL).
CC FT DOMAIN 61 70 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 71 91 2 (POTENTIAL).
CC FT DOMAIN 92 104 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 105 126 3 (POTENTIAL).
CC FT DOMAIN 127 143 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 144 168 4 (POTENTIAL).
CC FT DOMAIN 169 200 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 201 220 5 (POTENTIAL).

```

```

FT DOMAIN 221 237 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 238 262 6 (POTENTIAL).
FT DOMAIN 263 279 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 280 303 7 (POTENTIAL).
FT DOMAIN 304 354 CYTOPLASMIC (POTENTIAL).
FT DISULFID 103 180 BY SIMILARITY.
FT CARBOHYD 270 270 N-LINKED (GLCNAC..) (POTENTIAL).
SQ SEQUENCE 354 AA; 41030 MW; 77EDB368AA4C868D CRC64;

Query Match 69.1%; Score 775; DB 1; Length 354;
Best Local Similarity 78.2%; Pred. No. 5.5e-42;
Matches 147; Conservative 18; Mismatches 21; Indels 2; Gaps 1;

QY 1 MDQVSSP--TYDINYITSEPCQKINVKQAARLLPPLYSLVFIFGVGNMLVILINC 58
Db 1 MDFQGSIPTYIIDYDYSMSAPCKVNVKQIAAQLLPPLYSLVFIFGVGNMVFILISC 60
QY 59 KRKSMTDIYLLNLAISDLFFLLITVPFWAHYAAQWDFGNTMCLLGLYFGFSGIF 118
Db 61 KKLKSMTDIYLLNLAISDLFFLLITVPFWAHYAAQWDFGNTMCLLGLYFGFSGIF 120
QY 119 IILLTIDRYLAVHAFALKARTVTRGVTVITWVAVFASLPGLITRSQKGLHYTC 178
Db 121 IILLTIDRYLAVHAFALKARTVTRGVTVITWVAVFASLPGLITRSQKGLHYTC 180
QY 179 SSHPYIK 186
Db 181 SPHFLHIQ 188

```

Search completed: June 3, 2003, 19:22:56
Job time : 13.2338 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:20:46 ; Search time 23.5976 Seconds
(without alignments)
875.891 Million cell updates/sec

Title: US-09-938-703-6
Perfect score: 1122
Sequence: 1 MDQVSPYDINTYSEPC.....AACGHLLGNPKNSASVSK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 958 | 85.4 | 352 | A43113 | chemokine (C-C) re |
| 2 | 694 | 61.9 | 360 | JC2443 | chemokine (C-C) re |
| 3 | 694 | 61.9 | 374 | I38450 | chemokine (C-C) re |
| 4 | 598 | 53.3 | 355 | A45177 | chemokine (C-C) re |
| 5 | 573.5 | 51.1 | 355 | I49339 | macrophage inflamm |
| 6 | 547.5 | 48.8 | 360 | A57160 | chemokine (C-C) re |
| 7 | 546.5 | 48.7 | 360 | JC4587 | chemokine (C-C) re |
| 8 | 535.5 | 47.7 | 359 | I49341 | MIP-1 alpha recept |
| 9 | 506.5 | 45.1 | 355 | G02436 | chemokine (C-C) re |
| 10 | 489.5 | 43.6 | 356 | I49340 | MIP-1 alpha recept |
| 11 | 449 | 40.0 | 383 | S55594 | G protein-coupled |
| 12 | 427.5 | 38.1 | 355 | JC5067 | G protein-coupled |
| 13 | 386 | 34.4 | 354 | I58186 | probable G protein |
| 14 | 378 | 33.7 | 344 | JC5942 | chemokine receptor |
| 15 | 365 | 32.5 | 355 | JC4304 | orphan G protein-c |
| 16 | 345.5 | 30.8 | 378 | B55735 | lymphocyte-specifi |
| 17 | 342.5 | 30.5 | 378 | A55735 | G protein-coupled |
| 18 | 333.5 | 29.7 | 378 | A45680 | G protein-coupled |
| 19 | 331.5 | 29.5 | 352 | A45747 | neuropeptide Y/pep |
| 20 | 326 | 29.1 | 353 | S28787 | neuropeptide Y/pep |
| 21 | 325.5 | 29.0 | 352 | G00048 | fusin (LESTRA) - c |
| 22 | 321.5 | 28.7 | 359 | I51372 | angiotensin II rec |
| 23 | 320 | 28.5 | 359 | JC5068 | G protein-coupled |
| 24 | 317 | 28.3 | 369 | JC1194 | angiotensin II rec |
| 25 | 314 | 28.0 | 359 | A43656 | angiotensin II rec |
| 26 | 312 | 27.8 | 359 | JC1104 | angiotensin II rec |
| 27 | 312 | 27.8 | 359 | I39418 | angiotensin II rec |
| 28 | 311 | 27.7 | 359 | JC2134 | angiotensin II rec |
| 29 | 310 | 27.6 | 359 | S15403 | angiotensin II rec |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 30 | 309 | 27.5 | 359 | 2 | JH0621 | angiotensin II rec |
| 31 | 308 | 27.5 | 359 | 2 | A48857 | angiotensin II rec |
| 32 | 307 | 27.4 | 359 | 2 | S44425 | angiotensin II rec |
| 33 | 306 | 27.3 | 359 | 2 | JQ1516 | angiotensin II rec |
| 34 | 291 | 25.9 | 350 | 2 | JH0621 | G protein-coupled |
| 35 | 287.5 | 25.6 | 367 | 2 | JE0349 | interferon-inducib |
| 36 | 283 | 25.2 | 362 | 2 | JN0694 | angiotensin II rec |
| 37 | 276 | 24.6 | 333 | 2 | I65989 | G protein-coupled |
| 38 | 276 | 24.6 | 363 | 2 | I57955 | somatostatin recep |
| 39 | 276 | 24.6 | 364 | 2 | JN0763 | somatostatin recep |
| 40 | 274.5 | 24.5 | 359 | 2 | A48921 | interleukin-8 rece |
| 41 | 273.5 | 24.4 | 384 | 2 | A47249 | brain-specific som |
| 42 | 270.5 | 24.1 | 388 | 2 | JN0605 | somatostatin recep |
| 43 | 269 | 24.0 | 360 | 2 | A53611 | interleukin-8 rece |
| 44 | 267 | 23.8 | 363 | 2 | I37940 | somatostatin recep |
| 45 | 265 | 23.6 | 355 | 2 | JQ1231 | interleukin-8 rece |

ALIGNMENTS

RESULT 1

A43113
chemokine (C-C) receptor 5 - human
N:Alternate names: C-C CKR-5; CCR5
C:Species: Homo sapiens (Man)
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text change 20-Jun-2000
C:Accession: A43113; S71808; A58834; A58832; G02653; A58833
R:Samson, M.; Labbe, O.; Mollereau, C.; Vassart, G.; Parmentier, M.
Biochemistry 35, 3362-3367, 1996
A:Title: Molecular cloning and functional expression of a new human CC-chemokine rece
A:Reference number: A43113; MUID:96241590; PMID:8639485
A:Accession: A43113
A:Molecule type: mRNA
A:Residues: 1-352 <SAM>
A:Cross-references: GB:X91492; NID:g1262810; PIDN:CAA62796.1; PID:g1262811
R:Samson, M.; Libert, P.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Sara
M.; Imal, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;
Nature 382, 722-725, 1996
A:Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele
A:Reference number: S71808; MUID:96345670; PMID:8751444
A:Accession: S71808
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 182-206;207-230 <SAM2>
A:Accession: A58834
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-184; 'IKDSHLGAGPAACHGHLLGNPKNSASVSK' <SAM3>
A:Cross-references: GB:X99393; NID:g1524062; PIDN:CAA67767.1; PID:g1524063
A:Note: this frameshift mutation results in a non-functional receptor but confers a c
nd may have had a selective advantage by conferring resistance to Yersinia plague in;
R:Combiadere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.
J. Leukoc. Biol. 60, 147-152, 1996
A:Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine
A:Reference number: A58832; MUID:96295970; PMID:8699119
A:Accession: A58832
A:Molecule type: mRNA
A:Residues: 1-352 <COM1>
A:Cross-references: GB:U57840; NID:g1502408; PIDN:AAB17071.1; PID:g1502409
A:Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes
R:Combiadere, C.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01541
A:Accession: G02653
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-89; 'L' / 91-352 <COM2>
A:Cross-references: EMBL:U57840
R:Raport, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.
J. Biol. Chem. 271, 17161-17166, 1996
A:Title: Molecular cloning and functional characterization of a novel human CC chemol
A:Reference number: A58833; MUID:96291862; PMID:8663314

A;Accession: A58833
 A;Molecule type: mRNA
 A;Residues: 1-352 <RAP>
 A;Cross-references: GB:U54994; NID:g1457945; PIDN:AAC50598.1; PID:g1457946
 C;Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see PIR:A30575) and dual-tropic strains of HIV-1 bind to a complex of chemokine receptors.
 C;Genetics:
 A;Gene: GDB:CMKBR5; CCR5; CCR-5; CC-CR-5; CCR5; ChemR13
 A;Cross-references: GDB:1230510; OMIM:601373
 A;Map position: 3p21-3p21
 C;Function:
 A;Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES.
 A;Note: probably acts to control granulocyte proliferation and differentiation
 C;Superfamily: vertebrate rhodopsin
 C;Keywords: AIDs; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein
 F:32-56/Domain: transmembrane #status predicted <TM1>
 F:67-87/Domain: transmembrane #status predicted <TM2>
 F:103-124/Domain: transmembrane #status predicted <TM3>
 F:142-166/Domain: transmembrane #status predicted <TM4>
 F:193-218/Domain: transmembrane #status predicted <TM5>
 F:236-257/Domain: transmembrane #status predicted <TM6>
 F:285-300/Domain: transmembrane #status predicted <TM7>
 F:20-269,101-178/Disulfide bonds: #status predicted
 F:268/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted
 F:340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 85.4%; Score 958; DB 2; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.2e-80;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQSSPIYDINYYSEPCQKINVKQIAARLLPPLYSILVIFGVGNMVLILLINCKR 60
 DB 1 MDYQSSPIYDINYYSEPCQKINVKQIAARLLPPLYSILVIFGVGNMVLILLINCKR 60

QY 61 LKSMIDYLLNLAISDLFFLLVPPFWAHYAAQWDFGNMQLLTGLYFGFGIFGFI 120
 DB 61 LKSMIDYLLNLAISDLFFLLVPPFWAHYAAQWDFGNMQLLTGLYFGFGIFGFI 120

QY 121 LITDRYLAVHAFKARTVGVTVSVITWVAVFASLPGLIFTRSQKGLHYTCSS 180
 DB 121 LITDRYLAVHAFKARTVGVTVSVITWVAVFASLPGLIFTRSQKGLHYTCSS 180

QY 181 HEPY 184
 DB 181 HEPY 184

RESULT 2
 JC2443
 chemokine (C-C) receptor 2, splice form B - human
 N;Alternate names: C-C CR-2; monocyte chemoattractant protein 1 receptor; monocyte chemoattractant protein 1
 C;Species: Homo sapiens (man)
 C;Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 20-Jun-2000
 C;Accession: JC2443; I38463
 R;Yamagami, S.; Tokuda, Y.; Ishii, K.; Tanaka, H.; Endo, N.
 Biochem. Biophys. Res. Commun. 202, 1156-1162, 1994
 A;Title: cDNA cloning and functional expression of a human monocyte chemoattractant protein 1
 A;Reference number: JC2443; MUID:94324942; PMID:8048929
 A;Accession: JC2443
 A;Molecule type: mRNA
 A;Residues: 1-360 <YAM>
 A;Cross-references: DDBJ:D29984; NID:g531246; PIDN:BAA06253.1; PID:g531247
 R;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
 A;Title: Molecular cloning and functional expression of two monocyte chemoattractant proteins 1 and 2
 A;Reference number: A53477; MUID:94195821; PMID:8146186
 A;Accession: I38463
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-360 <RES>
 A;Cross-references: EMBL:U03905; NID:g472557; PIDN:AAA19120.1; PID:g472558
 C;Genetics:

A;Gene: GDB:CMKBR2
 A;Cross-references: GDB:337364; OMIM:601267
 A;Map position: 3p21-3p21
 C;Superfamily: vertebrate rhodopsin
 C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane protein
 F:43-70/Domain: transmembrane #status predicted <TM1>
 F:81-100/Domain: transmembrane #status predicted <TM2>
 F:115-136/Domain: transmembrane #status predicted <TM3>
 F:154-178/Domain: transmembrane #status predicted <TM4>
 F:207-226/Domain: transmembrane #status predicted <TM5>
 F:244-268/Domain: transmembrane #status predicted <TM6>
 F:287-309/Domain: transmembrane #status predicted <TM7>
 F:14/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:113-190/Disulfide bonds: #status predicted

Query Match 61.9%; Score 694; DB 2; Length 360;
 Best Local Similarity 75.9%; Pred. No. 2.4e-56;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYYTSEPCQKINVKQIAARLLPPLYSILVIFGVGNMVLILLINCKRLKSMIDYLL 69
 DB 24 FDYDY--GAPCHKFDVKQIGQALLPPLYSILVIFGVGNMVLILLINCKRLKCLDIYL 81

QY 70 LNLAIISDLFFLLVPPFWAHYAAQWDFGNMQLLTGLYFGFGIFGFIILLIDRYLA 129
 DB 82 LNLAIISDLFFLLVPPFWAHYAAQWDFGNMQLLTGLYFGFGIFGFIILLIDRYLA 141

QY 130 VYHAFKARTVGVTVSVITWVAVFASLPGLIFTRSQKGLHYTCSSHPF 183
 DB 142 IVHAFKARTVGVTVSVITWVAVFASLPGLIFTRSQKGLHYTCSSHPF 195

RESULT 3
 I38450
 chemokine (C-C) receptor 2, splice form A - human
 N;Alternate names: C-C CR-2; monocyte chemoattractant protein 1 receptor; monocyte chemoattractant protein 1
 C;Species: Homo sapiens (man)
 C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 13-Aug-1999
 C;Accession: I38450
 R;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 2752-2756, 1994
 A;Title: Molecular cloning and functional expression of two monocyte chemoattractant proteins 1 and 2
 A;Reference number: A53477; MUID:94195821; PMID:8146186
 A;Accession: I38450
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-374 <RES>
 A;Cross-references: EMBL:U03882; NID:g472555; PIDN:AAA19119.1; PID:g472556
 C;Genetics:

A;Gene: GDB:CMKBR2
 A;Cross-references: GDB:337364; OMIM:601267
 A;Map position: 3p21-3p21
 C;Superfamily: vertebrate rhodopsin
 C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane protein
 F:44-68/Domain: transmembrane #status predicted <TM1>
 F:79-99/Domain: transmembrane #status predicted <TM2>
 F:115-136/Domain: transmembrane #status predicted <TM3>
 F:154-178/Domain: transmembrane #status predicted <TM4>
 F:208-226/Domain: transmembrane #status predicted <TM5>
 F:244-265/Domain: transmembrane #status predicted <TM6>
 F:292-309/Domain: transmembrane #status predicted <TM7>
 F:14/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:32-277,113-190/Disulfide bonds: #status predicted

Query Match 61.9%; Score 694; DB 2; Length 374;
 Best Local Similarity 75.9%; Pred. No. 2.5e-56;
 Matches 132; Conservative 17; Mismatches 23; Indels 2; Gaps 1;

QY 10 YDINYYTSEPCQKINVKQIAARLLPPLYSILVIFGVGNMVLILLINCKRLKSMIDYLL 69
 DB 24 FDYDY--GAPCHKFDVKQIGQALLPPLYSILVIFGVGNMVLILLINCKRLKCLDIYL 81

QY 70 LNLAIISDLFFLLVPPFWAHYAAQWDFGNMQLLTGLYFGFGIFGFIILLIDRYLA 129

[illegible]

JC4304
 orphan G protein-coupled receptor - human
 N:Alternate names: V28 protein
 C:Species: Homo sapiens (man)
 C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 19-May-2000
 C:Accession: JC4304
 R:Report, C.J.; Schweickart, V.L.; Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
 Gene 163, 295-299, 1995
 A:Title: The orphan G-protein-coupled receptor-encoding gene V28 is closely related t
 A:Reference number: JC4304; MUID:96011651; PMID:7590284

C:Comment: This protein is a key regulator of many immune and homeostatic responses

C:Genetics:

A:Gene: V28

A:Map position: 3pter-p21

C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; lymphokine; transmembrane protein

F:35-57/Domain: transmembrane #status predicted <TM1>

F:66-88/Domain: transmembrane #status predicted <TM2>

F:104-125/Domain: transmembrane #status predicted <TM3>

F:146-165/Domain: transmembrane #status predicted <TM4>

F:197-217/Domain: transmembrane #status predicted <TM5>

F:230-254/Domain: transmembrane #status predicted <TM6>

F:275-296/Domain: transmembrane #status predicted <TM7>

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 32.5% | Score 365; | DB 2; | Length 355; |
| Best Local Similarity | 43.2%; | Pred. NO. 4.3e-26; | | |
| Matches 73; | Conservative 28; | Mismatches 64; | Indels 4; | Gaps 2; |

| matches | 73; | Conservative | 28; | mismatches | 04; | inlets | 4; | gaps | 2 |
|---------|-----|---|-----|------------|-----|--------|----|------|---|
| QY | 17 | SEPCQKINVRQIAARLLPPLYSLVIFGFGVMYVILLINCKRKSKMTDYLVLNLAISD | 76 | | | | | | |
| | : | : | : | : | : | : | : | : | : |
| Db | 18 | AEACIGDIVVGVNVLSTFTSVFALGIVGLNVLVVFALNKKSKPSTVDYLINLAISD | 77 | | | | | | |

| | | | |
|----|-----|---|-----|
| QY | 77 | LEFLLTVPFNAHVAARAQDNGNTWCQLLGLYIGFGSGFEFFILITLTDRLAVHVAFA | 136 |
| Db | 78 | LEFVATLDFPHTYLINKEGLHNAKCKTAAFFIGFGSFIFFITVIDRILAVIAAANS | 137 |
| QY | 137 | LKARTVTGGVTVSYIVWVAVFAFLPGLIFTIRSQEKGELHUTCSSHPYII | 185 |
| Db | 138 | MNNRTVQGVTSISLGVWAAAIIIVAAPQFMFTK-QEN---ECLGDYVEP | 182 |

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:17:26 ; Search time 7.93902 Seconds
(without alignments)
520.312 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDHLGAGPAAACHGHLHLLGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A.Geneseq.101002.*
- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | ID | Description |
|------------|-------------|-------|--------|----|-----------------------------|
| 1 | 164 | 100.0 | 215 | 18 | AAW27408 Inactive human CCR |
| 2 | 164 | 100.0 | 215 | 20 | AAW88238 HIV-1 co-receptor |
| 3 | 55 | 33.5 | 84 | 22 | ABB36928 Peptide #4504 enco |
| 4 | 55 | 33.5 | 84 | 22 | AAW57726 Human brain expres |
| 5 | 55 | 33.5 | 84 | 22 | AAW70138 Human bone marrow |
| 6 | 55 | 33.5 | 84 | 23 | ABG39777 Human peptide enco |
| 7 | 55 | 33.5 | 94 | 22 | ABG23873 Novel human diagno |
| 8 | 54 | 32.9 | 276 | 22 | ABB69618 Drosophila melanog |
| 9 | 53.5 | 32.6 | 189 | 22 | ABG27770 Novel human diagno |
| 10 | 53.5 | 32.6 | 197 | 22 | ABG27246 Novel human diagno |

| | | | | | |
|----|------|------|------|----|------------------------------|
| 11 | 52.5 | 32.0 | 216 | 21 | AAG18166 Arabidopsis thalia |
| 12 | 52.5 | 32.0 | 265 | 21 | AAG18165 Arabidopsis thalia |
| 13 | 52.5 | 32.0 | 270 | 21 | AAAG28063 Arabidopsis thalia |
| 14 | 52.5 | 32.0 | 478 | 21 | AAAG28062 Arabidopsis thalia |
| 15 | 51 | 31.1 | 193 | 22 | ABG21033 Novel human diagno |
| 16 | 49 | 29.9 | 278 | 21 | AAAG22079 Arabidopsis thalia |
| 17 | 49 | 29.9 | 278 | 21 | AAAG43094 Arabidopsis thalia |
| 18 | 49 | 29.9 | 290 | 21 | AAAG22078 Arabidopsis thalia |
| 19 | 49 | 29.9 | 290 | 21 | AAAG43093 Arabidopsis thalia |
| 20 | 49 | 29.9 | 331 | 21 | AAAG22077 Arabidopsis thalia |
| 21 | 49 | 29.9 | 331 | 21 | AAAG43092 Arabidopsis thalia |
| 22 | 48.5 | 29.6 | 698 | 22 | ABB59005 Drosophila melanog |
| 23 | 48 | 29.3 | 101 | 22 | AAO07177 Human polypeptide |
| 24 | 48 | 29.3 | 186 | 23 | ABP42972 Human ovarian anti |
| 25 | 48 | 29.3 | 297 | 23 | AAU77123 Human protease #2. |
| 26 | 48 | 29.3 | 320 | 22 | AAU19238 Human G protein-co |
| 27 | 48 | 29.3 | 340 | 22 | AAU40849 Propionibacterium |
| 28 | 48 | 29.3 | 451 | 23 | AAU77122 Human protease #1. |
| 29 | 48 | 29.3 | 486 | 23 | AAU77124 Human protease #3. |
| 30 | 48 | 29.3 | 828 | 22 | ABB63141 Drosophila melanog |
| 31 | 48 | 29.3 | 1186 | 23 | AAU72895 Human metalloprote |
| 32 | 48 | 29.3 | 1189 | 23 | AAU74750 Human protease PRT |
| 33 | 48 | 29.3 | 1210 | 23 | AAU85415 Human protein NOV1 |
| 34 | 48 | 29.3 | 1213 | 23 | AAU77128 Human protease #7. |
| 35 | 48 | 29.3 | 1216 | 23 | AAU77127 Human protease #6. |
| 36 | 48 | 29.3 | 1219 | 23 | AAU77126 Human protease #5. |
| 37 | 48 | 29.3 | 1222 | 23 | AAU77125 Human protease #4. |
| 38 | 48 | 29.3 | 1223 | 23 | ABG30863 Human metalloprote |
| 39 | 48 | 29.3 | 1223 | 23 | AAU79747 Human ADAM-TS 1 pr |
| 40 | 48 | 29.3 | 1232 | 23 | AAU77130 Human protease #9. |
| 41 | 48 | 29.3 | 1235 | 23 | AAU77129 Human protease #8. |
| 42 | 48 | 29.3 | 1249 | 23 | AAU77132 Human protease #11 |
| 43 | 48 | 29.3 | 1252 | 23 | AAU77131 Human protease #10 |
| 44 | 47.5 | 29.0 | 146 | 22 | ABG19311 Novel human diagno |
| 45 | 47.5 | 29.0 | 971 | 22 | ABB63091 Drosophila melanog |

ALIGNMENTS

RESULT 1
AAW27408
ID AAW27408 standard; Protein; 215 AA.
XX AC AAW27408;
DT 14-APR-1998 (first entry)
DE Inactive human CCR5.
XX
KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
XX
OS Homo sapiens.
XX
PN WO9732019-A2.
XX
PD 04-SEP-1997.
XX
PF 28-FEB-1997; 97WO-BE00023.
XX
PR 06-AUG-1996; 96EP-0870102.
PR 01-MAR-1996; 96EP-0870021.
XX
PA (EURO-) EUROSCREEN SA.
XX
PI Libert F, Parmentier M, Samson M, Vassart G;
XX WPI; 1997-479829/44.
DR

XX N-PSDB; AAT90118.
 PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 PT useful to diagnose, prevent and/or treat inflammatory disorders,
 PT autoimmune disease and viral infection
 XX
 PS Claim 7; Fig 1d-e; 94pp; English.
 XX
 CC The present sequence is an inactive human CC (Cys-Cys)
 CC chemokine receptor 5 (CCR5), which lacks the last 3 transmembrane
 CC regions and the regions involved in G protein-coupling. CCR5 or
 CC its cDNA can be used to diagnose, treat and/or prevent inflammatory
 CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,
 CC idiopathic pulmonary fibrosis and psoriasis, viral infections,
 CC especially human immunodeficiency virus type 1 or type 2 (HIV-1 or
 CC HIV-2) infection, cancer, atherosclerosis and autoimmune disorders.
 CC Subjects that express the inactive receptor have a predisposition,
 CC or resistance to HIV-1 and/or HIV-2.
 XX
 SQ Sequence 215 AA;
 Query Match 100.0%; Score 164; DB 18; Length 215;
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 31
 |||||
 DB 185 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 215
 |||||
 RESULT 2
 AAW8238
 ID AAW8238 standard; Protein; 215 AA.
 XX
 AC AAW8238;
 XX
 DT 15-MAR-1999 (first entry)
 XX
 DE HIV-1 co-receptor CCR5 variant CCR5-delta32.
 XX
 KW HIV-1; CCR5; CCR5-delta32; co-receptor; infection; diagnosis; AIDS;
 KW gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 32..56
 FT /note= "transmembrane domain 1"
 FT Domain 67..87
 FT /note= "transmembrane domain 2"
 FT Domain 103..124
 FT /note= "transmembrane domain 3"
 FT Domain 142..167
 FT /note= "transmembrane domain 4"
 XX
 PN WO9854317-AL.
 XX
 PD 03-DEC-1998.
 XX
 PF 29-MAY-1998; 98WO-EP03437.
 XX
 PR 30-MAY-1997; 97US-0048057.
 XX
 PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
 XX
 PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
 XX
 DR WPI; 1999-059835/05.
 DR N-PSDB; AAW84159.
 XX
 PT New CCR5 variant protein of the HIV-1 co-receptor - useful in
 PT developing resistance of CCR5-expressing cells to HIV-1 infection
 XX

PS Disclosure; Page 38-39; 55pp; English.
 XX
 CC This is the amino acid sequence of a CCR5 variant protein,
 CC designated CCR5-delta32, that includes the first 4 transmembrane
 CC domains of wild-type CCR5 (see AAW88232), but lacks transmembrane
 CC domains 5-7. CCR5 serves as a co-receptor for infection by
 CC macrophage-tropic (M-tropic) strains of HIV-1. Individuals
 CC homozygous for the CCR5-delta32 mutation are resistant to HIV-1
 CC infection, but heterozygous individuals are susceptible. The
 CC invention additionally relates to the identification of variant
 CC CCR5s (see AAW88231), which lacks transmembrane domains 3-7 of
 CC CCR5. The detection of CCR5 variants may be used to identify
 CC individuals at lower risk of infection relative to the general
 CC population who, if infected, may exhibit slower progression to
 CC AIDS. Probes and primers (see AAW84127-36) are provided for use in
 CC diagnostic methods for detecting the presence of such variants. A
 CC method is provided for inhibiting HIV-1 infection of a cell
 CC expressing the CCR5 receptor. This involves introducing a nucleic
 CC acid encoding a CCR5 variant into the cell, thereby reducing the
 CC number of functional CCR5 molecules present on the cell surface.
 XX
 SQ Sequence 215 AA;
 Query Match 100.0%; Score 164; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 31
 |||||
 DB 185 IKDSLHGAGPAAACHGHLILGNPKNSASVSK 215
 |||||
 RESULT 3
 ABB36998
 ID ABB36998 standard; Peptide; 84 AA.
 XX
 AC ABB36998;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #4504 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 FI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 FS Claim 27; SEQ ID NO 29633; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84 AA;

Query Match 33.5%; Score 55; DB 22; Length 84;
Best Local Similarity 50.0%; Pred. No. 3.3;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26
||||| |||||
Db 48 DSHLGAGPAAACHGHLILGNPKNS 65

RESULT 4
AAM57726
ID AAM57726 standard; Protein; 84 AA.

XX AAM57726;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29831.

XX Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX Example 4; SEQ ID NO: 29831; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.

XX Sequence 84 AA;

Query Match 33.5%; Score 55; DB 22; Length 84;
Best Local Similarity 50.0%; Pred. No. 3.3;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26
||||| |||||
Db 48 DSHLGAGPAAACHGHLILGNPKNS 65

RESULT 5
AAM70138
ID AAM70138 standard; Protein; 84 AA.

XX AAM70138;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30444.

XX Human: bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 30444; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.

XX Sequence 84 AA;

Query Match 33.5%; Score 55; DB 22; Length 84;
Best Local Similarity 50.0%; Pred. No. 3.3;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY 3 DSHLGAGPAAACHGHLILGNPKNS 26
||||| |||||
Db 48 DSHLGAGPAAACHGHLILGNPKNS 65

RESULT 6
ABG39777
ID ABG39777 standard; Peptide; 84 AA.

XX ABG39777;

XX 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 29442.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX Homo sapiens.
OS
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
PT
XX
XX Claim 27; SEQ ID NO 29442; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein

CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84 AA;
Query Match 33.5%; Score 55; DB 23; Length 84;
Best Local Similarity 50.0%; Pred. NO. 3.3;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;
QY 3 DSHLGAGPAACGHHLLGNPKNS 26
||||| ||||| | : |
Db 48 DSHLGAGPAATA-----GGPRTS 65
RESULT 7
ABG23873
ID ABG23873 standard; Protein; 94 AA.
XX
XX AC ABG23873;
XX
XX DT 18-FEB-2002 (first entry)
XX
XX DE Novel human diagnostic protein #23864.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS88060.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID NO 54232; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences, (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

db 167 GAGPAAACI AVCCSPGSSHHHIGHVGHITATGHP 199

RESULT 11
AAG18166

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.0%; Score 52.5; DB 21; Length 216;
Best Local Similarity 45.5%; Pred. No. 22;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDSHLGAGPAACHGHLGNP 23
::: ||| |||||
Db 84 RNDNGYGGPGG-HGHLWAGNP 104

RESULT 12
AAG18165
ID AAG18165 standard; Protein; 265 AA.

XX AAG18165;

AC AAG18165;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 19465.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

| | | | | |
|----|--------------|---------------|----|------------------------------|
| PR | 31-AUG-1999; | 99US-0151438. | PN | EP1033405-A2. |
| PR | 01-SEP-1999; | 99US-0151930. | XX | |
| PR | 07-SEP-1999; | 99US-0152363. | PD | 06-SEP-2000. |
| PR | 10-SEP-1999; | 99US-0153070. | XX | |
| PR | 13-SEP-1999; | 99US-0153758. | PF | 25-FEB-2000; 2000EP-0301439. |
| PR | 15-SEP-1999; | 99US-0154018. | XX | |
| PR | 16-SEP-1999; | 99US-0154039. | PR | 25-FEB-1999; 99US-0121825. |
| PR | 20-SEP-1999; | 99US-0154779. | PR | 03-MAR-1999; 99US-0123180. |
| PR | 22-SEP-1999; | 99US-0155139. | PR | 09-MAR-1999; 99US-0123548. |
| PR | 23-SEP-1999; | 99US-0155486. | PR | 23-MAR-1999; 99US-0125788. |
| PR | 24-SEP-1999; | 99US-0155659. | PR | 25-MAR-1999; 99US-0126264. |
| PR | 28-SEP-1999; | 99US-0156458. | PR | 29-MAR-1999; 99US-0126785. |
| PR | 29-SEP-1999; | 99US-0156596. | PR | 01-APR-1999; 99US-0127462. |
| PR | 04-OCT-1999; | 99US-0157117. | PR | 06-APR-1999; 99US-0128234. |
| PR | 05-OCT-1999; | 99US-0157753. | PR | 08-APR-1999; 99US-0128714. |
| PR | 06-OCT-1999; | 99US-0157865. | PR | 16-APR-1999; 99US-0129845. |
| PR | 07-OCT-1999; | 99US-0158029. | PR | 19-APR-1999; 99US-0130077. |
| PR | 08-OCT-1999; | 99US-0158232. | PR | 21-APR-1999; 99US-0130449. |
| PR | 12-OCT-1999; | 99US-0158369. | PR | 23-APR-1999; 99US-0130510. |
| PR | 13-OCT-1999; | 99US-0159293. | PR | 23-APR-1999; 99US-0130891. |
| PR | 13-OCT-1999; | 99US-0159294. | PR | 28-APR-1999; 99US-0131449. |
| PR | 13-OCT-1999; | 99US-0159295. | PR | 30-APR-1999; 99US-0132048. |
| PR | 14-OCT-1999; | 99US-0159329. | PR | 30-APR-1999; 99US-0132407. |
| PR | 14-OCT-1999; | 99US-0159330. | PR | 04-MAY-1999; 99US-0132484. |
| PR | 14-OCT-1999; | 99US-0159331. | PR | 05-MAY-1999; 99US-0132485. |
| PR | 14-OCT-1999; | 99US-0159637. | PR | 06-MAY-1999; 99US-0132486. |
| PR | 14-OCT-1999; | 99US-0159638. | PR | 06-MAY-1999; 99US-0132487. |
| PR | 18-OCT-1999; | 99US-0159584. | PR | 07-MAY-1999; 99US-0132863. |
| PR | 21-OCT-1999; | 99US-0160741. | PR | 11-MAY-1999; 99US-0134256. |
| PR | 21-OCT-1999; | 99US-0160767. | PR | 14-MAY-1999; 99US-0134218. |
| PR | 21-OCT-1999; | 99US-0160768. | PR | 14-MAY-1999; 99US-0134219. |
| PR | 21-OCT-1999; | 99US-0160770. | PR | 14-MAY-1999; 99US-0134221. |
| PR | 21-OCT-1999; | 99US-0160814. | PR | 14-MAY-1999; 99US-0134370. |
| PR | 21-OCT-1999; | 99US-0160815. | PR | 18-MAY-1999; 99US-0134768. |
| PR | 22-OCT-1999; | 99US-0160980. | PR | 19-MAY-1999; 99US-0134941. |
| PR | 22-OCT-1999; | 99US-0160981. | PR | 20-MAY-1999; 99US-0135124. |
| PR | 22-OCT-1999; | 99US-0160989. | PR | 21-MAY-1999; 99US-0135353. |
| PR | 25-OCT-1999; | 99US-0161404. | PR | 24-MAY-1999; 99US-0135629. |
| PR | 25-OCT-1999; | 99US-0161405. | PR | 25-MAY-1999; 99US-0136021. |
| PR | 25-OCT-1999; | 99US-0161406. | PR | 27-MAY-1999; 99US-0136392. |
| PR | 26-OCT-1999; | 99US-0161359. | PR | 28-MAY-1999; 99US-0136782. |
| PR | 26-OCT-1999; | 99US-0161360. | PR | 01-JUN-1999; 99US-0137222. |
| PR | 26-OCT-1999; | 99US-0161361. | PR | 03-JUN-1999; 99US-0137528. |
| PR | 28-OCT-1999; | 99US-0161920. | PR | 04-JUN-1999; 99US-0137502. |
| PR | 28-OCT-1999; | 99US-0161992. | PR | 07-JUN-1999; 99US-0137724. |
| PR | 28-OCT-1999; | 99US-0161993. | PR | 08-JUN-1999; 99US-0138094. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 10-JUN-1999; 99US-0138540. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 10-JUN-1999; 99US-0138847. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 14-JUN-1999; 99US-0139119. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 16-JUN-1999; 99US-0139452. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 16-JUN-1999; 99US-0139453. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 17-JUN-1999; 99US-0139492. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139454. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139455. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139456. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139457. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139458. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139459. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139460. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139461. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139462. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139463. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139750. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139763. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 21-JUN-1999; 99US-0139817. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 22-JUN-1999; 99US-0139899. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 23-JUN-1999; 99US-0140353. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 23-JUN-1999; 99US-0140354. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 24-JUN-1999; 99US-0140695. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 24-JUN-1999; 99US-0140823. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 29-JUN-1999; 99US-0140991. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 30-JUN-1999; 99US-0141287. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 01-JUL-1999; 99US-0141842. |

Query Match 32.0%; Score 52.5; DB 21; Length 265;
Best Local Similarity 45.5%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
QY 2 KDSHLGAGPAACHGHLILGNP 23
Db 133 RNDNGYCPGG-HGLMAGNP 153
RESULT 13
AAG28063
ID AAG28063 standard; Protein; 270 AA.
XX
AC AAG28063;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33140.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144088.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.0%; Score 52.5; DB 21; Length 270;
Best Local Similarity 45.5%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDSHLGAGPAAACHGHLGNP 23
::: ||| |||:
Db 138 RNDNNGYPPGG-HGHLGNP 158

RESULT 14

AAG28062
ID AAG28062 standard; Protein; 478 AA.

XX AC AAG28062;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 33139.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-01493175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 51392; 103pp; English.
PS
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

```

AA      Sequence      193 AA;
Query Match      31.1%;      Score 51;      DB 22;      Length 193;
Best Local Similarity 45.8%;      Pred. No. 31;
Matches 11;      Conservative 1;      Mismatches 6;      Indels 6;      Gaps 1

QY      3      DSHLAGGPAACGCHLLIGNPKNS 26
          ||||| |||
          ||:|
          ||:|
Db       55      DSHLAGGPACTA-----GGPRTS 72

Search completed: June 3, 2003, 19:22:35
Job time : 8.93902 sec

```

Job time : 8.93902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:24:01 ; Search time 51.565 Seconds
(without alignments)
422.051 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122
Sequence: 1 MDQVSSPIYDINITYSEPC.....AACHGHLGNPKNSASYK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 303519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/FCI_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 1122 | 100.0 | 215 | 10 US-09-938-719-6 | Sequence 6, Appli |
| 2 | 1122 | 100.0 | 215 | 10 US-09-939-226-6 | Sequence 6, Appli |
| 3 | 1122 | 100.0 | 215 | 10 US-09-938-703-6 | Sequence 6, Appli |
| 4 | 958 | 85.4 | 184 | 10 US-09-938-719-4 | Sequence 4, Appli |
| 5 | 958 | 85.4 | 184 | 10 US-09-939-226-4 | Sequence 4, Appli |
| 6 | 958 | 85.4 | 184 | 10 US-09-938-703-4 | Sequence 4, Appli |
| 7 | 958 | 85.4 | 352 | 9 US-10-232-686-2 | Sequence 2, Appli |
| 8 | 958 | 85.4 | 352 | 9 US-10-086-814-1 | Sequence 1, Appli |
| 9 | 958 | 85.4 | 352 | 9 US-09-734-221A-14 | Sequence 14, Appli |
| 10 | 958 | 85.4 | 352 | 10 US-09-725-285-2 | Sequence 2, Appli |
| 11 | 958 | 85.4 | 352 | 10 US-09-759-841-2 | Sequence 2, Appli |
| 12 | 958 | 85.4 | 352 | 10 US-09-779-879A-22 | Sequence 22, Appli |
| 13 | 958 | 85.4 | 352 | 10 US-09-779-880A-22 | Sequence 22, Appli |
| 14 | 958 | 85.4 | 352 | 10 US-09-813-653-15 | Sequence 15, Appli |
| 15 | 958 | 85.4 | 352 | 10 US-09-796-202-1 | Sequence 1, Appli |
| 16 | 958 | 85.4 | 352 | 10 US-09-195-662A-2 | Sequence 2, Appli |
| 17 | 958 | 85.4 | 352 | 10 US-09-339-912A-2 | Sequence 2, Appli |
| 18 | 958 | 85.4 | 352 | 10 US-09-938-719-5 | Sequence 5, Appli |
| 19 | 958 | 85.4 | 352 | 10 US-09-939-226-5 | Sequence 5, Appli |

```
20 958 85.4 352 10 US-09-938-703-5 Sequence 5, Appli
21 958 85.4 352 10 US-09-502-783A-2 Sequence 2, Appli
22 958 85.4 352 12 US-10-106-623-2 Sequence 2, Appli
23 952 84.8 352 10 US-09-813-653-17 Sequence 17, Appli
24 943 84.0 352 10 US-09-779-879A-2 Sequence 2, Appli
25 943 84.0 352 10 US-09-779-880A-2 Sequence 2, Appli
26 933 83.2 352 12 US-10-106-623-20 Sequence 20, Appli
27 695 61.9 360 10 US-09-131-827A-20 Sequence 20, Appli
28 694 61.9 344 9 US-10-232-686-9 Sequence 9, Appli
29 694 61.9 344 10 US-09-779-879A-9 Sequence 9, Appli
30 694 61.9 344 10 US-09-779-880A-9 Sequence 9, Appli
31 694 61.9 347 10 US-09-104-792-3 Sequence 3, Appli
32 694 61.9 360 10 US-09-131-827A-2 Sequence 2, Appli
33 692 61.7 360 10 US-09-938-719-7 Sequence 7, Appli
34 692 61.7 360 10 US-09-939-226-7 Sequence 7, Appli
35 692 61.7 360 10 US-09-938-703-7 Sequence 7, Appli
36 608.5 54.2 329 10 US-09-725-285-9 Sequence 9, Appli
37 608.5 54.2 329 10 US-09-195-662A-9 Sequence 9, Appli
38 608.5 54.2 329 10 US-09-339-912A-9 Sequence 9, Appli
39 608.5 54.2 329 10 US-09-502-783A-9 Sequence 9, Appli
40 598 53.3 355 10 US-09-961-068-1 Sequence 1, Appli
41 598 53.3 355 10 US-09-960-547-1 Sequence 1, Appli
42 598 53.3 375 9 US-10-219-834-78 Sequence 78, Appli
43 594 52.9 355 10 US-09-938-719-9 Sequence 9, Appli
44 594 52.9 355 10 US-09-939-226-9 Sequence 9, Appli
45 594 52.9 355 10 US-09-938-703-9 Sequence 9, Appli
```

ALIGNMENTS

```
RESULT 1
US-09-938-719-6
; Sequence 6, Application US/09938719
; Patent No. US20020106742A1
GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION NUMBER: US/09/938,719
APPLICATION NUMBER: US/09/938,719
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
```

US-09-938-719-6

Query Match 100.0%; Score 1122; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLFLLTPFVAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLFLLTPFVAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFII 120
QY 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPYIKDHLGAGPAAACHGHLILGNPKNSASVSK 215
DB 181 HPYIKDHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 2

US-09-939-226-6
; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-226-6

Query Match 100.0%; Score 1122; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60

DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLFLLTPFVAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLFLLTPFVAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFII 120
QY 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
DB 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPYIKDHLGAGPAAACHGHLILGNPKNSASVSK 215
DB 181 HPYIKDHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 3

US-09-938-703-6
; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,703
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-703-6

Query Match 100.0%; Score 1122; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
DB 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVFIQFVGNMVLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFLFLLTPFVAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFII 120
DB 61 LKSMTDIYLLNLAIISDLFLFLLTPFVAHYAAQWDFGNTMCOLLTGLYFIFGFSGIFII 120
QY 121 LITDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180

Db 121 LTTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPYKDSHLGAGPAAACHGHLILGNPKNSASVSK 215
Db 181 HFPYKDSHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 4

US-09-938-719-4
; Sequence 4, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,719
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-938-719-4

Query Match 85.4%; Score 958; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 7.5e-88;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFFLLVFPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFII 120
Db 61 LKSMTDIYLLNLAISDLFFLLVFPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFII 120
QY 121 LTTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LTTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 5

US-09-939-226-4
; Sequence 4, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-939-226-4
Query Match 85.4%; Score 958; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 7.5e-88;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAISDLFFLLVFPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFII 120
Db 61 LKSMTDIYLLNLAISDLFFLLVFPFWAHYAAQWDFGNTMCQLLTGLYIFGFFSGIFII 120
QY 121 LTTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LTTIDRYLAVHVAFAKARTVFGVTVSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184
RESULT 6
US-09-938-703-4
; Sequence 4, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK

RESULT 9
US-09-734-221A-14
; Sequence 14, Application US/09734221A
; Publication No. US20030096221A1

```

; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William C.
; TITLE OF INVENTION: SUBSTITUTED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 61010-AB-1
; CURRENT APPLICATION NUMBER: US/10/086,814
; CURRENT FILING DATE: 2002-02-28
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-086-814-1

Query Match      85.4%; Score 958; DB 9; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1 MDYQSSPIYDINYTISEPCQINKVKQIARLLPPLYSIVFIFGVGNMLVILILINCKR 60
        |||
Db      1 MDYQSSPIYDINYTISEPCQINKVKQIARLLPPLYSIVFIFGVGNMLVILILINCKR 60

Qy      61 LKSMTDIYLLNLAISDLFLLTPVPWHAHYAAQWDFGNTMCQLTGLYFGFSGIFPII 120
        |||
Db      61 LKSMTDIYLLNLAISDLFLLTPVPWHAHYAAQWDFGNTMCQLTGLYFGFSGIFPII 120

Qy      121 LLRIDRYLVAVHVFALKARTVTFGVVTSVITWVAVFASLPGLIITRSQKEGLHYTCSS 180
        |||
Db      121 LLRIDRYLVAVHVFALKARTVTFGVVTSVITWVAVFASLPGLIITRSQKEGLHYTCSS 180

Qy      181 HFPY 184
        ||||
Db      181 HFPY 184

RESULT 9
US-09-734-221A-14
; Sequence 14, Application US/09734221A
; Publication No. US20030096221A1

```

```

RESULT 10
US-09-725-285-2
; Sequence 2, Application US/09725285
; Patent No. US2001000241A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGNR10
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/725,285
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence: Genomic
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-725-285-2

Query Match      85.4%; Score 958; DB 10; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MDYQSSPTIDNYITSEPCQKINVKQIAARLLPPLYSIVLTFGVGNMLVILLINCKR 60
DB      1 MDYQSSPTIDNYITSEPCQKINVKQIAARLLPPLYSIVLTFGVGNMLVILLINCKR 60

QY      61 LKSMTDIYLLMAISDLFELLVPPWAIHYAAQWDGNTMQLTGLYFTGFSGIFPII 120
DB      61 LKSMTDIYLLMAISDLFELLVPPWAIHYAAQWDGNTMQLTGLYFTGFSGIFPII 120

QY      121 LTIIDRYLVAVHYAFALKARTVFGVTSVITWWVAFASLPGLIFTRSKESGLHYTCSS 180
DB      121 LTIIDRYLVAVHYAFALKARTVFGVTSVITWWVAFASLPGLIFTRSKESGLHYTCSS 180

QY      181 HPFY 184
DB      181 HPFY 184

RESULT 11
US-09-759-841-2
; Sequence 2, Application US/09759841
; Patent No. US20010039026A1
; GENERAL INFORMATION:
; APPLICANT: Rickett, Graham A
; APPLICANT: Dobbs, Susan
; APPLICANT: Perros, Manoussos
; TITLE OF INVENTION: Assay Method
; FILE REFERENCE: PCI0348PME
; CURRENT APPLICATION NUMBER: US/09/759,841
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: GB 0000661.9
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: GB 0000663.5
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: GB 0000659.3
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens

```


US-09-779-880A-22

Query Match 85.4%; Score 958; DB 10; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCOKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
|||||
Db 1 MDYQVSSPIYDINYYTSEPCOKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
|||||

QY 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYIFGFGSIFPII 120
|||||
Db 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYIFGFGSIFPII 120
|||||

QY 121 LTTDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
|||||
Db 121 LTTDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
|||||

QY 181 HFPY 184
|||||
Db 181 HFPY 184

RESULT 15
US-09-796-202-1
; Sequence 1, Application US/09796202
; Patent No. US20020068813a1
; GENERAL INFORMATION:
; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 2048/61010/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/796,202
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: human
US-09-796-202-1

Query Match 85.4%; Score 958; DB 10; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCOKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
|||||
Db 1 MDYQVSSPIYDINYYTSEPCOKINVKQIAARLLPPLYSLVIFGFGVGNMLVILLINCKR 60
|||||

QY 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYIFGFGSIFPII 120
|||||
Db 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYIFGFGSIFPII 120
|||||

QY 121 LTTDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
|||||
Db 121 LTTDRYLAVHVAFAKARTVGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
|||||

QY 181 HFPY 184
|||||
Db 181 HFPY 184

Search completed: June 3, 2003, 19:35:59
Job time : 51.565 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:21:31 ; Search time 4.53659 Seconds
(without alignments)
201.057 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/pdata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/pdata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/pdata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/pdata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/pdata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/pdata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|-------------------|-------------------|
| 1 | 164 | 100.0 | 34 | US-08-833-752-13 | Sequence 13, Appl |
| 2 | 164 | 100.0 | 215 | US-09-087-232A-17 | Sequence 17, Appl |
| 3 | 164 | 100.0 | 215 | US-08-833-752-6 | Sequence 6, Appl |
| 4 | 50 | 30.5 | 501 | US-08-660-963-13 | Sequence 13, Appl |
| 5 | 47.5 | 29.0 | 2304 | US-09-324-867-4 | Sequence 4, Appl |
| 6 | 47.5 | 29.0 | 2319 | US-08-212-133A-8 | Sequence 8, Appl |
| 7 | 47.5 | 29.0 | 2319 | US-08-474-503-6 | Sequence 6, Appl |
| 8 | 47.5 | 29.0 | 2319 | US-08-670-707A-6 | Sequence 6, Appl |
| 9 | 47.5 | 29.0 | 2319 | US-09-037-601-6 | Sequence 6, Appl |
| 10 | 47.5 | 29.0 | 2319 | US-09-315-179-6 | Sequence 6, Appl |
| 11 | 47.5 | 29.0 | 2319 | US-09-523-656-28 | Sequence 28, Appl |
| 12 | 47.5 | 29.0 | 2319 | PCT-US94-13200-6 | Sequence 6, Appl |
| 13 | 45.5 | 27.7 | 1438 | US-09-209-316-1 | Sequence 1, Appl |
| 14 | 45.5 | 27.7 | 1471 | US-08-683-839B-3 | Sequence 3, Appl |
| 15 | 45.5 | 27.7 | 1661 | US-08-882-083-2 | Sequence 2, Appl |
| 16 | 45.5 | 27.7 | 1661 | US-08-558-107-2 | Sequence 2, Appl |
| 17 | 45.5 | 27.7 | 1661 | US-09-243-539-2 | Sequence 2, Appl |
| 18 | 45.5 | 27.7 | 1958 | US-07-945-283-2 | Sequence 2, Appl |
| 19 | 45.5 | 27.7 | 2332 | US-07-864-004B-4 | Sequence 4, Appl |
| 20 | 45.5 | 27.7 | 2332 | US-08-251-937A-4 | Sequence 4, Appl |
| 21 | 45.5 | 27.7 | 2332 | US-08-212-133A-2 | Sequence 2, Appl |
| 22 | 45.5 | 27.7 | 2332 | US-08-276-594A-2 | Sequence 2, Appl |
| 23 | 45.5 | 27.7 | 2332 | US-08-474-503-2 | Sequence 2, Appl |
| 24 | 45.5 | 27.7 | 2332 | US-08-670-707A-2 | Sequence 2, Appl |
| 25 | 45.5 | 27.7 | 2332 | US-09-037-601-2 | Sequence 2, Appl |
| 26 | 45.5 | 27.7 | 2332 | US-09-324-867-3 | Sequence 3, Appl |
| 27 | 45.5 | 27.7 | 2332 | US-09-315-179-2 | Sequence 2, Appl |

28 45.5 27.7 2332 4 US-09-523-656-2 Sequence 2, Appl
29 45.5 27.7 2332 5 PCT-US93-03275-4 Sequence 4, Appl
30 45.5 27.7 2332 5 PCT-US94-13200-2 Sequence 2, Appl
31 45.5 27.7 2351 1 US-08-121-202-2 Sequence 2, Appl
32 45.5 27.7 2351 1 US-08-366-851A-2 Sequence 2, Appl
33 45.5 27.7 2351 6 5171844-2 Patent No. 5171844
34 45.5 27.7 2351 6 5422260-1 Patent No. 5422260
35 45 27.4 272 4 US-09-177-165A-27 Sequence 27, Appl
36 45 27.4 815 4 US-09-177-165A-24 Sequence 24, Appl
37 44.5 27.1 166 3 US-08-513-974B-312 Sequence 312, Appl
38 44.5 27.1 2232 4 US-09-091-219-25 Sequence 25, Appl
39 44.5 27.1 2247 4 US-09-091-219-2 Sequence 2, Appl
40 44 26.8 188 2 US-08-933-750C-8 Sequence 8, Appl
41 44 26.8 188 4 US-09-234-613-8 Sequence 8, Appl
42 44 26.8 455 2 US-08-870-827-3 Sequence 3, Appl
43 44 26.8 455 4 US-09-317-179-3 Sequence 32, Appl
44 43.5 26.5 782 4 US-09-543-084A-32 Sequence 33, Appl
45 43.5 26.5 782 4 US-09-543-084A-33 Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-08-833-752-13
; Sequence 13, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-833-752-13

Query Match 100.0%; Score 164; DB 4; Length 34;
Best Local Similarity 100.0%; Pred. No. 8.9e-18;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31
|||||
Db 4 IKDHLGAGPAAACHGILLGNPKNSASVSK 34

RESULT 2

```
US-09-087-232A-17
; Sequence 17, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-087-232A-17

Query Match 100.0%; Score 164; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASYSK 31
Db 185 IKDHLGAGPAAACHGHLILGNPKNSASYSK 215

RESULT 3
US-08-833-752-6
; Sequence 6, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-6

Query Match 100.0%; Score 164; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 7.6e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASYSK 31
Db 185 IKDHLGAGPAAACHGHLILGNPKNSASYSK 215

RESULT 4
US-08-660-963-13
; Sequence 13, Application US/08660963
; Patent No. 5852187
; GENERAL INFORMATION:
; APPLICANT: Thorne, Michael O.
; APPLICANT: Gaylinn, Bruce D.
; APPLICANT: Horikawa, Reiko
; APPLICANT: Lyons Jr., Charles E.
; TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
; TITLE OF INVENTION: GROWTH HORMONE RELEASING HORMONE RECEPTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,963
; FILING DATE: 12-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Shaughnessy, Brian P.
; REGISTRATION NUMBER: 32,747
; REFERENCE/DOCKET NUMBER: 18046.036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-660-963-13
```

SEQUENCE CHARACTERISTICS:

LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO N-terminal
FRAGMENT TYPE: N-terminal
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Lakich, D.
AUTHORS: Gitschler, J.
TITLE: Sequence of the Murine Factor VIII cDNA.
Patent No. 574446
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
US-08-474-503-6

Query Match 29.0%; Score 47.5; DB 1; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY 2 KDSHLG-AGPAAACHGHLNPKNSASVS 30
DB 1832 RDMHSLGILLCANTL--NPAHGRQVS 1859

RESULT 8

US-08-670-707A-6
Sequence 6, Application US/08670707A
Patent No. 5859204
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 2319 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Mus musculus
PUBLICATION INFORMATION:
AUTHORS: Elder, F.
AUTHORS: Lakich, D.
AUTHORS: Gitschler, J.
TITLE: Sequence of the Murine Factor VIII cDNA
Patent No. 5859204
JOURNAL: Genomics
VOLUME: 16
PAGES: 374-379
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
US-08-670-707A-6

Query Match 29.0%; Score 47.5; DB 2; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY 2 KDSHLG-AGPAAACHGHLNPKNSASVS 30
DB 1832 RDMHSLGILLCANTL--NPAHGRQVS 1859

RESULT 9

US-09-037-601-6
Sequence 6, Application US/09037601
Patent No. 6180371
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089


```

; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.
; AUTHORS: Lachich, D.
; AUTHORS: Gitschier, J.
; TITLE: Sequence of the Murine Factor VIII cDNA.
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 2319
PCT-US94-13200-6

```

```

Query Match      29.0%; Score 47.5; DB 5; Length 2319;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

```

```

QY      2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db      1832 RDMHSLGLGPLLICHANTL--NPAHGRQVS 1859

```

```

RESULT 13
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-tuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

```

```

Query Match      27.7%; Score 45.5; DB 4; Length 1438;
Best Local Similarity 40.0%; Pred. No. 2.6e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

```

```

QY      2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db      951 KDVSGLGLPLLVCHTNTL--NPAHGRQVT 978

```

```

RESULT 14
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
; GENERAL INFORMATION:
; APPLICANT: Ill, Charles . R. et al.
; TITLE OF INVENTION: Use of Viral Cis-Acting Post-Transcriptional
; TITLE OF INVENTION: Regulatory Sequences to Increase Expression of
; TITLE OF INVENTION: Introns Genes Containing Near-Consensus Splice Sites
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,839B
; FILING DATE: 11-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: TTI-138
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1471 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-839B-3

```

```

Query Match      27.7%; Score 45.5; DB 1; Length 1471;
Best Local Similarity 40.0%; Pred. No. 2.7e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

```

```

QY      2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db      984 KDVSGLGLPLLVCHTNTL--NPAHGRQVT 1011

```

```

RESULT 15
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5859292
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,083
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1661 amino acids
; TYPE: amino acid

```


; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-882-083-2

Query Match 27.7%; Score 45.5; DB 2; Length 1661;
Best Local Similarity 40.0%; Pred. No. 3.1e+02;
Matches 12; Conservative 3; Mismatches 12; Indels 3; Gaps 2;

QY 2 KDSHLG-AGPAAACHGHLILGNPKNSASVS 30
Db 1174 KDVHSLGLIGPLLYCHTNTL--NPAHGQVT 1201

Search completed: June 3, 2003, 19:25:11
Job time : 5.53659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:44:09 ; Search time 27.6947 Seconds
(without alignments)
3063.973 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2628
Sequence: 1 GAATTCGCCACAGAGCA.....AGTAGATGATCGGAATTC 1442

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09938703/runat_03062003_161408.22885/app_query.fasta_1.1870
-DB=Issued Patents AA -OFFT=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPTCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938703.ecgn_1.1.42 -runat_03062003_161408.22885 -NCPD=6 -ICPD=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIAG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*
1: /cgn2.6/ptodata/1/iaa/5A.COMB.pcp.*
2: /cgn2.6/ptodata/1/iaa/5B.COMB.pcp.*
3: /cgn2.6/ptodata/1/iaa/6A.COMB.pcp.*
4: /cgn2.6/ptodata/1/iaa/6B.COMB.pcp.*
5: /cgn2.6/ptodata/1/iaa/PTCUS.COMB.pcp.*
6: /cgn2.6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 1745.5 | 66.4 | 352 | 4 | US-09-087-232A-13 |
| 2 | 1745.5 | 66.4 | 352 | 4 | US-08-861-105-14 |
| 3 | 1745.5 | 66.4 | 352 | 4 | US-08-575-967A-2 |
| 4 | 1745.5 | 66.4 | 352 | 4 | US-08-833-752-5 |
| 5 | 1739.5 | 66.2 | 352 | 4 | US-09-045-583-52 |
| 6 | 1739.5 | 66.2 | 352 | 4 | US-09-534-185-52 |
| 7 | 1730.5 | 65.8 | 352 | 3 | US-08-466-343D-2 |
| 8 | 1718.5 | 65.4 | 352 | 4 | US-09-517-605-5 |
| 9 | 1463.5 | 55.7 | 354 | 4 | US-08-724-984A-2 |
| 10 | 1331 | 50.6 | 347 | 1 | US-08-461-244-3 |
| 11 | 1331 | 50.6 | 360 | 1 | US-08-450-393A-4 |
| 12 | 1331 | 50.6 | 360 | 4 | US-08-446-669-4 |

| | | | | | | |
|----|--------|------|-----|---|-------------------|-------------------|
| 13 | 1331 | 50.6 | 360 | 4 | US-09-045-583-50 | Sequence 50, Appl |
| 14 | 1331 | 50.6 | 360 | 4 | US-09-534-185-50 | Sequence 50, Appl |
| 15 | 1331 | 50.6 | 360 | 5 | PCT-US95-00476-4 | Sequence 4, Appl |
| 16 | 1317 | 50.1 | 360 | 4 | US-08-833-752-7 | Sequence 7, Appl |
| 17 | 1311.5 | 49.9 | 360 | 4 | US-09-045-583-51 | Sequence 51, Appl |
| 18 | 1311.5 | 49.9 | 360 | 4 | US-09-534-185-51 | Sequence 51, Appl |
| 19 | 1191 | 45.3 | 374 | 1 | US-08-450-393A-2 | Sequence 2, Appl |
| 20 | 1191 | 45.3 | 374 | 4 | US-08-446-669-2 | Sequence 2, Appl |
| 21 | 1191 | 45.3 | 374 | 5 | PCT-US95-00476-2 | Sequence 9, Appl |
| 22 | 1190.5 | 45.3 | 344 | 3 | US-08-466-343D-9 | Sequence 9, Appl |
| 23 | 1122 | 42.7 | 215 | 4 | US-09-087-232A-17 | Sequence 17, Appl |
| 24 | 1122 | 42.7 | 215 | 4 | US-08-833-752-6 | Sequence 6, Appl |
| 25 | 1003.5 | 38.2 | 355 | 1 | US-08-012-988A-2 | Sequence 2, Appl |
| 26 | 1003.5 | 38.2 | 355 | 1 | US-08-450-393A-5 | Sequence 5, Appl |
| 27 | 1003.5 | 38.2 | 355 | 4 | US-08-446-669-5 | Sequence 5, Appl |
| 28 | 1003.5 | 38.2 | 355 | 4 | US-09-239-938-1 | Sequence 1, Appl |
| 29 | 1003.5 | 38.2 | 355 | 5 | PCT-US95-00476-5 | Sequence 5, Appl |
| 30 | 976.5 | 37.2 | 355 | 4 | US-08-833-752-9 | Sequence 9, Appl |
| 31 | 958 | 36.5 | 184 | 4 | US-08-833-752-4 | Sequence 4, Appl |
| 32 | 949.5 | 36.1 | 355 | 4 | US-09-045-583-53 | Sequence 53, Appl |
| 33 | 949.5 | 36.1 | 355 | 4 | US-09-534-185-53 | Sequence 4, Appl |
| 34 | 912 | 34.7 | 355 | 4 | US-08-575-967A-4 | Sequence 4, Appl |
| 35 | 912 | 34.7 | 355 | 4 | US-08-847-296B-1 | Sequence 1, Appl |
| 36 | 912 | 34.7 | 355 | 4 | US-09-045-583-54 | Sequence 54, Appl |
| 37 | 912 | 34.7 | 355 | 4 | US-09-534-185-54 | Sequence 54, Appl |
| 38 | 891.5 | 33.9 | 360 | 4 | US-08-875-573-20 | Sequence 20, Appl |
| 39 | 891.5 | 33.9 | 360 | 4 | US-09-232-878-2 | Sequence 2, Appl |
| 40 | 891.5 | 33.9 | 360 | 4 | US-09-045-583-55 | Sequence 55, Appl |
| 41 | 891.5 | 33.9 | 360 | 4 | US-09-534-185-55 | Sequence 55, Appl |
| 42 | 873 | 33.2 | 355 | 4 | US-08-833-752-8 | Sequence 8, Appl |
| 43 | 857.5 | 32.6 | 360 | 4 | US-08-833-752-10 | Sequence 10, Appl |
| 44 | 716 | 27.2 | 355 | 1 | US-08-461-244-2 | Sequence 2, Appl |
| 45 | 716 | 27.2 | 355 | 4 | US-09-045-583-56 | Sequence 56, Appl |

ALIGNMENTS

RESULT 1
US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quilient et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087/232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
 LENGTH: 352 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-087-232A-13

Alignment Scores:

Pred. No.: 5,68e-182 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-087-232A-13 (1-352)

```

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGCAGCCCGCTCTCGCTCGCTCTACTCTACTGTGTG 359
DB 21 GlnYsIleAsnValIysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTGTGTTTGTGGCCACATGCTGGTCATCCCTCACTCATCTGATTAACATGCAAGG 419
DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysIysArg 60
QY 420 CTGAAGAGCATGATCATCTACTCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
DB 61 LeuYsSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTACTCTGCCCTCTGGGCTCACTATGCTGGCCGCGCCAGTGGAGCTTTGGAAATACATG 539
DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCACATCTGTACAGGCTCTATTATATAGGCTTCTCTGGAATCTCTTCATCATC 599
DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIle 120
QY 600 CTCTGCAATCAATGATGATCTGCTGGCTGCTGCTCAATGCTGTGTTTAAAGCCAGG 659
DB 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGTCACTTGGGCTGTCACAGTGTGATCACTGGTGGTGGTGGTGGTGGTGGTGGTCT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTTACAGATCTCAAAAGAGCTTTCATTTACACCTGCAGCTCT 779
DB 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnLysGlnLysTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGAGATGATCAT 806
DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTGGGCTGGTCTCGCGCTGCTGTGTCATGCTCATCTGCTACTCGGGAATCTTAAAC 866
DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTGGTGTGCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
DB 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValAlaArgLeuIlePheThr 240
QY 927 CATGATGTTTATTCTTCTTCGGGCTCCCTACACATGCTCTTCCTTCCTGCAACCTT 986
DB 240 eMeIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTGGCTGAATATTCAGTAGTCTTAACAGGTTGAGCCAGCTATGCA 1046
DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280

```

RESULT 2

US-08-861-105-14
 Sequence 14, Application US/08861105
 Patent No. 6258527

GENERAL INFORMATION:

APPLICANT: LITTMAN, DAN R.
 APPLICANT: DENG, HONGKUI
 APPLICANT: ELLMEIER, WILFRIED
 APPLICANT: LANDAU, NATHANIEL R.
 APPLICANT: LIU, RONG
 TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
 TITLE OF INVENTION: MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
 TITLE OF INVENTION: USES THEREOF
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.
 STREET: 411 Hackensack Ave, Continental Plaza, 4th
 STREET: Floor
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/861,105

FILING DATE:

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/666,020

FILING DATE: 19-JUN-1996

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/227,319

FILING DATE: 13-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 1049-1-004 N1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 352 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: Homo sapiens

US-08-861-105-14

Alignment Scores:

Pred. No.: 5,68e-182 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-861-105-14 (1-352)

QY 240 ATGGATTATCAAGTCTCAAGTCAATCTATGACATCAATATTATATACATCGGACCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGCAGCCGCCCTCCCTGCTCCCTACTACTACACGGTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATTAACAGTCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGACATGACATCACTACCTGCTCAACCTGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTACTGTCCTCTCGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATPACAAATG 539
Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaLagInTrpAspPheGlyAsnThrMet 100
QY 540 TGTCACTCTTACAGAGGCTCTATTATTAGGCTCTCTCTGGAATCTCTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATCAGTAGTACCTGCTGCTGCCATGCTGTTTGTCTTTAAAGCCAGG 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTACCTTTGGGGTGGGACAGTGTATCATCTGGTGGTGGCTGCTGTTGGCTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCTGAGATCATCTTTACAGATCTCAAAAGAGTCTTCTATTACACCTGACGTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTGGGGCTGGCTCGGCTGCTGCTCATGCTCATCTGCTACTCGGAATCCTTAAAC 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTCTGAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTCAACAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrII 240
QY 927 CATGATTTGTTATTTTCTCTCTGCTGCTCCCTCAACATGCTCTCTCTCAACACCTT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGAACTCTTGGCTGGAATATTGCAGTAGCTCTAACAGGTTCGACCAAGTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGACTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAGTCTCAGAACTACCTTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320

QY 1167 CTCGAAATGCTGTCTATTATTTCCACGAGAGTCCCGAGGAGCAGTCAAGTCAATTAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnIleSerValGlyLeu 352

RESULT 3

US-08-575-967A-2
; Sequence 2, Application US/08575967A
; Patent No. 6265184
; GENERAL INFORMATION:
; APPLICANT: Gray et al.
; TITLE OF INVENTION: Chemokine Receptor Materials and Methods
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/575,967A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6265184and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32918
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-485-1900
; TELEFAX: 206-485-1662
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: /- "88C amino acid sequence"
US-08-575-967A-2

Alignment Scores:
Pred. No.: 5,68e-182 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-575-967A-2 (1-352)

QY 240 ATGGATTATCAAGTCTCAAGTCAATCTATGACATCAATATTATATACATCGGACCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAATCGCAGCCGCCCTCCCTGCTCCCTACTACTACCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATTAACAGTCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60

QY 420 CTGAAGACGATGACATGACATGCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTCTACTGCTCCCTTCTGGCTCACTATGCTGCCCGCCAGTGGACCTTTGGAAATACAAATG 539
 Db 81 LeuThrValProPheThrPalahisTyrAlaAlaGlnTrpaspPheGlyAsnThrMet 100
 QY 540 TGTCACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCTGACATCGATAGTACCTGGCTGCTGCTCCATGCTGCTGCTTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCTACCTTTGGGTGGTGCACAGTGTGATCACTGGGTGGTGGTGTGGTCTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTACAGATCTCAAAAGAGGTCTTCTTACCTGACGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
 QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValII 200
 QY 807 CTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 Db 220 rLeuLeuArgCysArgGlnGluLysLysArgHisArgAlaValArgLeuIlePheThrII 240
 QY 927 CATGATGTTTATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMeIleValTyrPheLeuPheThrPalahisTyrAsnLeuLeuLeuLeuLeuLeuLeu 260
 QY 987 CAGGAATCTTTGGCTGGAATATGACAGTCTTACAGAGTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetGI 280
 QY 1047 GTCACAGACTCTGGGATGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGGAGAGTTCAGAACTACCTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGGAATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATTCACCTGGGAGGAGAAATATCTGCTGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 4

US-08-833-752-5

; Sequence 5, Application US/08833752

; Patent No. 6448375

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; APPLICANT: PARMENTIER, MARC

; APPLICANT: VASSART, GILBERT

; APPLICANT: LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/833,752
 FILING DATE: 9-APR-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E

REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER:

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS:

LENGTH: 352 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-833-752-5

Alignment Scores:

Pred. No.: 5 68e-182 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-833-752-5 (1-352)

QY 240 ATGGAATATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATACATGAGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
 QY 300 CAAAATAATCAATGTAAGCAAAATCGACCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTyrSerLeuVal 40
 QY 360 TTATCTTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATAAACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGACGATGACATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
 QY 480 CTCTACTGCTCCCTTCTGGGCTCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 81 LeuThrValProPheThrPalahisTyrAlaAlaGlnTrpaspPheGlyAsnThrMet 100
 QY 540 TGTCACTCTTGACAGGCTCTATTTATAGGCTTCTTCTTCTGGAATCTTCTTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCTCTGACATTCATAGGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCTACCTTTGGGTGGTGCACAGTGTGATCACTGGGTGGTGGTGTGGTCTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTACAGATCTCAAAAGAGGTCTTCTTACCTGACGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180

```

QY 780 CATTTCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTTGGGCTGGCTGGCGCTGCTGCTCATGCTATCTGCTACTCGGGAATCCTAAAC 866
Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTGTGAAATGAGAGAGAGGACAGGCGTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGTTTATTTCTCTCTGCGCTCCCTACAACTGCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGTAATATGACAGTGTCTACAGGTTGGACCAAGTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGACATCTTGGATGAGCAGCAGTGTGATCAACCCCATCTATGCTTTGT 1106
Db 280 nValThrGlnThrLeuGlyMetThrHisCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAGAGTTGAGAACTACTCTTACTCTCTTCTCAAAAGCACATTGCCAAGCCT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTTCTTATTTCCAGCAAGAGCTCCGACGAGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCATCGGGAGCAGGAATATCTCTGGGCTTG 1263
Db 340 rArgSerThrGlyGlnGlnLysSerValGlyLeu 352

RESULT 5
US-09-045-583-52
; Sequence 52, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)42-4214
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-045-583-52

Alignment Scores:
Pred. No.: 2,56e-181 Length: 352
Score: 1739.50 Matches: 338
Percent Similarity: 96.60% Conservative: 3
Best Local Similarity: 95.75% Mismatches: 0
Query Match: 66.19% Indels: 12
DB: 4 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-045-583-52 (1-352)
QY 240 ATGGATATCAAGTGTCAAGTCCATCTATGACATCAATATTATATCATCGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAspTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAAAATGCCAGCCCGCTCTCTCGCTCCGCTCTACTACTGGTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTCTGGGCAACATGCTGGTCACTCTCTCTCTCTCTCTCTCTCTCTCT 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGACATGATGATGATCTACCTGCTCAAGCTGGCCATCTCTGACCTGTGTTTCTCT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTCTCTGGGCTCAGTATGCTGCGCGCCAGCTGGGACTTGGAAATCAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTCTGGAATCTCTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATGATAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 AGGTCACCTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTTACAGATCTCAAAAAGAGGTCTTCTTACCTACCTGACGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTTGGGCTGGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTGTGAAATGAGAGAGGACAGGCGTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGTTTATTTCTCTCTGCGCTCCCTACAACTGCTCTCTCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGTAATATGACAGTGTCTACAGGTTGGACCAAGTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGACATCTTGGATGAGCAGCAGTGTGATCAACCCCATCTATGCTTTGT 1106

```


Patent No. 6025154
 GENERAL INFORMATION:
 APPLICANT: LI, YI
 TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
 TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGNR10 (AS AMENDED)
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., NW, SUITE 600
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/466,343D
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 352 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-466-343D-2

Alignment Scores:
 Pred. No.: 2,46e-180 Length: 352
 Score: 1730.50 Matches: 336
 Percent Similarity: 96.32% Conservative: 4
 Best Local Similarity: 95.18% Mismatches: 1
 Query Match: 63.85% Indels: 12
 DB: 3 Gaps: 1

US-09-938-703-3 (1-1442) x US-08-466-343D-2 (1-352)

QY 240 ATGATATCAAGTCAAGTCCAAATCTATGACATCAATATTATATACATCGGACCGCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAATCAATGTAAGCAAAATCGACGCCGCTCTGCTCCGCTCTACTCTACTGCTG 359
 Db 21 ProLysIleAsnValGlyGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTGGTGGCAACATCTGTGTCATCTCTCATCTCTGATTAATGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysGlnArg 60
 QY 420 CTGAAGAGCATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuGluSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTCTACTGCTCTCTGCTGCTCATCTGCTGCCGCGGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 81 LeuThrValProPheTrpAlaIleIleTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCCTGCAATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTCACCTTTGGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCTGAGTAATCATCTTACCAAGTCTCTCAAAAGAGCTCTTCAATACACCTGCTGCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTATCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
 QY 807 CTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTTGTTTATTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
 QY 987 CCAGGAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetGl 280
 QY 1047 GGTGACAGACTCTTGGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGACATTCGCAAAAGCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTTCTATTTCACAGAGAGGCTCCCGAGGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGGAAATATCTGTGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluLeuSerValGlyLeu 352

RESULT 8
 US-09-517-605-5
 ; Sequence 5, Application US/09517605
 ; Patent No. 6391567
 ; GENERAL INFORMATION:
 ; APPLICANT: Littman, Dan R.
 ; APPLICANT: Kwon, Douglas S.
 ; APPLICANT: van Kooyk, Yvette
 ; APPLICANT: Geijtenbeck, Theo
 ; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
 ; TITLE OF INVENTION: CELLS
 ; FILE REFERENCE: 1049-1-017
 ; CURRENT APPLICATION NUMBER: US/09/517,605
 ; CURRENT FILING DATE: 2000-03-02
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-517-605-5

Alignment Scores:
 Pred. No.: 5,02e-179 Length: 352
 Score: 1718.50 Matches: 334
 Percent Similarity: 95.75% Conservative: 4
 Best Local Similarity: 94.62% Mismatches: 3

| | | | |
|--|--------|---|------|
| Query Match: | 65.39% | Indels: | 12 |
| DB: | 4 | Gaps: | 1 |
| US-09-938-703-3 (1-1442) x US-09-517-605-5 (1-352) | | | |
| QY | 240 | ATGGATTATCAAGTGCATCAATCTATGACATCAATATTATATACATCGGAGCCCTGC | 299 |
| DB | 1 | MetAspPheGlnValSerProThrTyAspPheAspTyAspThrSerGluProCys | 20 |
| QY | 300 | CAAAAATCAATGTGAAGCAAAATCGAGCCCGCTCGCTCGCTCTACTACTCTGCTG | 359 |
| DB | 21 | GlnLysLeuValLysGlnLeuAlaAlaArgLeuLeuProLeuTySerLeuVal | 40 |
| QY | 360 | TTCACTTTGGTTTGTGGCAACATGCTGGTCACTCATCTCACTCACTCACTCACT | 419 |
| DB | 41 | PheIlePheGlyPheValGlyAsnMetLeuValLeuValLeuLeuAsnCysLysArg | 60 |
| QY | 420 | CTCAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT | 479 |
| DB | 61 | LeuLysSerMetThrAspPheLeuLysLeuLeuValLeuValLeuValLeuValLeu | 80 |
| QY | 480 | CTTACTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 539 |
| DB | 81 | LeuThrValProPheThrAlaHisTyAlaAlaGlnTrpAspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCACACTCTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTCTTCATCATC | 599 |
| DB | 101 | CysGlnLeuLeuThrGlyLeuTyPheIleGlyPheSerGlyLeuPhePheIleIle | 120 |
| QY | 600 | CTCCTGACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG | 659 |
| DB | 121 | LeuLeuThrIleAspArgTyLeuAlaIleValHisAlaValPheAlaLeuLysAla | 140 |
| QY | 660 | ACGGTCACCTTTGGGGTGGTGAAGTGTGATCACTTGGTGGTGGTGGTGGTGGT | 719 |
| DB | 141 | ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAla | 160 |
| QY | 720 | CTCCAGGAATCATCTTACCAAGATCTCAAAAGAGAGTCTCATACCTGCGAGCT | 779 |
| DB | 161 | LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyThrCysSer | 180 |
| QY | 780 | CATTTPCCATAC-----ATTAAAGATAGTCAT | 806 |
| DB | 181 | HisPheProTySerGlnTyGlnPheThrLysAsnPheGlnThrLeu-LysIleVal | 200 |
| QY | 807 | CTGGGGCTGGTCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 866 |
| DB | 200 | eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTySerGlyIleLeuLysTh | 220 |
| QY | 867 | TCGTCTCGGTGTCGAATGACAGAGACAGGCTGTGAGGCTTATCTTCACTCATC | 926 |
| DB | 220 | rLeuLeuArgCysArgAsnGlnLysArgGHisArgAlaValArgLeuIlePheThr | 240 |
| QY | 927 | CATGATGTTTATTTCTCTCTGCGGCTCCCTACACATTTGCTCTCTGCAACACCT | 986 |
| DB | 240 | eMetIleValTyPheLeuPheThrAlaProTyAsnIleValLeuLeuLeuAsnThrPh | 260 |
| QY | 987 | CCAGGAATCTTTGGCTGGAATATGTCAGTAGCTCTACACAGGTGGACCAAGCTAT | 1046 |
| DB | 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer | 280 |
| QY | 1047 | GGTGACAGAGACTCTGGGATGACGACTGCTGCATCAACCCCATCATCTATGCTT | 1106 |
| DB | 280 | nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPhe | 300 |
| QY | 1107 | CGGGGAGAAGTTCAGAACTTACTCTTATGCTTCTTCCAAAGACAGATTTGCCA | 1166 |
| DB | 300 | IGlyGluLysPheArgAsnTyLeuLeuValPhePheGlnLysHisIleAlaLysHis | 320 |
| QY | 1167 | CTGCAATGCTGTTCTATTTTCCAGAGAGGCTCCCGAGGACGACAGCTCAGTTT | 1226 |
| DB | 320 | eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyThr | 340 |

```
Db 81 PheLeuLeuThrLeuProPheThrPalahisYrAlaAlaSnGluTrpIlePheGlyAsn 100
QY 534 ACAATGTGCAACTCTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTCTTC 593
Db 101 IleMetCysLysValPheThrGlyValYrHisIleGlyYrPheGlyGlyLePhePhe 120
QY 594 ATCATCTCTGCAACATGATAGTACTGCTCTGCTCCATGCTGCTGCTGCTTTTAAA 653
Db 121 IleIleLeuLeuThrIleAspArgYrLeuAlaIleValHisAlaValPheAlaLeuLys 140
QY 654 GCCAGAGCGCTCACTTGGGTGGTGCAGAGTGTGATCATGCTGGTGGTGGTGTGTT 713
Db 141 ValThrValAsnPheGlyValIleThrSerValValThrIrpValValAlaValPhe 160
QY 714 GCGTCTCTCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTTTCATTCACCTGC 773
Db 161 AlaserLeuProGluIleIlePheThrArgSerGlnLysGluGlyPheHisYrThrCys 180
QY 774 AGCTCTCATTTCCATAC-----ATTAAAGAT 800
Db 181 SerProHisPheProHisThrGlnYrHisPheTrpLysSerPheGlnThrLeu-LysMe 200
QY 801 AGTCATCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 860
Db 200 tValIleLeuSerLeuLeuLeuLeuProLeuLeuValMetIleIleCysYrSerGlyIle 220
QY 861 AAAAATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 920
Db 220 uHisThrLeuPheArgCysArgAspGluLysArgHisArgAlaValArgLeuIlePh 240
QY 921 CACCATCATGATGTTATTTCTCTCTGGGCTCCCTACACATGTCCTCTCTCTGAA 980
Db 240 eAlaIleMetIleValYrPheLeuPheThrPheProTyraHisIleValLeuLeuLeuTh 260
QY 981 CACCTTCCAGGAATCTTGGCTGGAATATGTCAGTGTCTCAAGAGTGGACCAAGC 1040
Db 260 rThrPheGlnGluPhePheGlyLeuAsnAspCysSerSerSerAsnArgLeuAspGlnAl 280
QY 1041 TATGAGGTGACAGACTCTGGGATGACGACTGTGTCATGCTGCTCAACCCATCATCATGC 1100
Db 280 aMetGlnAlaThrGluThrLeuGlyMetThrHisCysCysLeuAsnProValIleYrAl 300
QY 1101 CTTTGTGCGGAGAGTTCAGAACTACTCTTATGCTTCTTCCAAAGACACATTCGCA 1160
Db 300 aPheValGlyGluLysPheArgSerYrLeuSerValPhePheArgYrHisIleVally 320
QY 1161 ACCTTCTGCAATGCTGTTCTATTTTCCAGAGAGGCTCCCGAGCGAGCAAGCTCAGT 1220
Db 320 sArgPheCysLysArgCysSerIlePheGlnGlnAspAsnProAspArgValSerVa 340
QY 1221 TTACACCGATCCACTGGGAGCAGGAAATATCTGGGCTG 1263
Db 340 lYrThrArgSerThrGlyGluHisGluValSerThrGlyLeu 354
```

RESULT 10

US-08-461-244-3

; Sequence 3, Application US/08461244

; Patent No. 5776729

; GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel R.

; APPLICANT: Yi, Li

; APPLICANT: Ruben, Steven M.

; APPLICANT: Rosen, Craig A.

; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,

; ADDRESS: STUART & OLSTEIN

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: USA

```
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,244
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fattaro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-445
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-244-3
```

Alignment Scores:

```
Pred. No.: 9,81e-137 Length: 347
Score: 1331.00 Matches: 258
Percent Similarity: 83.76% Conservative: 36
Best Local Similarity: 73.50% Mismatches: 46
Query Match: 50.65% Indels: 11
Db: 1 Gaps: 3
```

US-09-938-703-3 (1-1442) x US-08-461-244-3 (1-347)

```
QY 234 AACAGATGATTATCAAGTGTCAAGTCAATCTATGACATCAATATTATATACATCGGAG 293
Db 1 AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla 17
QY 294 CCCTGCCCAAAATCAATGTGAAGCAAAATCGCAGCCGCTCTCTGCTGCTGCTTACTCA 353
Db 18 ProCysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuYrSer 37
QY 354 CTGGTGTTCATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 413
Db 38 LeuValPheIlePheGlyPheValGlyAsnMetLeuValLeuLeuLeuLeuLeuLeu 57
QY 414 AAAAGGCTGAAGAGCATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTT 473
Db 58 LysLysLeuLysCysLeuThrAspIleYrLeuLeuAsnLeuAlaIleSerAspLeuLeu 77
QY 474 TTCCTTCTTACTGTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
Db 78 PheLeuIleThrLeuProLeuTrpAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn 97
QY 534 ACAATGTGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCTGGAATCTTCTTC 593
Db 98 AlaMetCysLysLeuPheThrGlyLeuYrHisIleGlyYrPheGlyGlyIlePhePhe 117
QY 594 ATCATCTCTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653
Db 118 IleIleLeuLeuThrIleAspArgYrYrLeuAlaIleValHisAlaValPheAlaLeuLys 137
QY 654 GCCAGGAGGTCACCTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 713
Db 138 AlaArgThrValThrPheGlyValValThrSerValIleThrIrpLeuValAlaValPhe 157
QY 714 GCGTCTCTCCAGGATCATCTTTACCATCTCAAAAGAGAGTCTTTCATTCACCTGC 773
Db 158 AlaSerValProGlyIleIlePheThrThrLysCysGlnLysGluAspSerValYrValCys 177
QY 774 AGC-----TCTCATTTTCCATATCAATTAAGATAGTATCTTTGGG 812
```


; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,583
FILING DATE: 20-MAR-98
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/POCKET NUMBER: MNI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 360 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-09-045-583-50

Alignment Scores:

Pred. No.: 9.96e-137 Length: 360
Score: 1331.00 Matches: 258
Percent Similarity: 83.76% Conservative: 36
Best Local Similarity: 73.50% Mismatches: 46
Query Match: 50.65% Indels: 11
Gaps: 3

US-09-938-703-3 (1-1442) x US-09-045-583-50 (1-360)

QY 234 AACAGATGATTTATCAAGTCTCAAGTCCAAATCTATGACATCAATATTATACATCGGAG 293
DB 14 AAGGluSerGlyGluGluValThr---PhePheAspTyrAspTyr-----GlyAla 30
QY 294 CCCTGCCAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCTCTCGCTCCGCTACTCA 353
DB 31 ProcysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer 50
QY 354 CTGCTGTTCATCTTTGGTTGGGCAACATGCTGGTCATCTCTCATCTGATTAACATGC 413
DB 51 LeuValPheIlePheGlyPheValGlyAsnMetLeuValValLeuLeuLeuLeuLeuCys 70
QY 414 AAAAGGCTGAAGACATGACATCTACCTGCTCAACCTGCGCATCTCTGACCTGTTT 473
DB 71 LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeu 90
QY 474 TTCCTCTTACTGTCCTCTCTGGCTCACTATGTCGCCGCCAGCGGACGCTTTGGAAT 533
DB 91 PheLeuIleThrLeuProLeuTyrAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn 110
QY 534 ACATGTGTCACTCTACAGGCTCTATTATTAGCTCTCTCTCTGGAATCTTCTTC 593
DB 111 AlaMetCysLeuLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyIlePhePhe 130
QY 594 ATCATCTCTCTGACATGATGATGCTGGCTGCTGCTGCTGCTGCTGCTGCTTAA 653
DB 131 IleIleLeuLeuThrIleAspTyrLeuAlaIleValHisAlaValPheAlaLeuLys 150
QY 654 GCCAGGACGCTACCTTTGGGTTGGTGCACAAAGTGTATCATCTGGTGGTGGCTGTGTTT 713

DB 151 AlaArgThrValThrPheGlyValValThrSerValThrTrpLeuValAlaValPhe 170
QY 714 GGTCTCTCCAGGAATCATCTTTACCAGATCTCAAAAGAGGTCTTCTTACACCTGC 773
DB 171 AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys 190
QY 774 AGC-----TCTCATTTTCCATACATTAAGATAGTCATCTCTGGG 812
DB 191 GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuG 210
QY 813 GCTGTCTCTCCGCTGCTGTCTCATCTCTACTCTCGGGAATCTCTAAACCTCTGCT 872
DB 210 yLeuValLeuProLeuIleMetValIleCysTyrSerGlyIleLeuLysThrIleLe 230
QY 873 TCGGTGTGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCCACCATCATCAT 932
DB 230 uArgCysArgAsnGluLysArgHisArgAlaValArgValIlePheThrIleMetI 250
QY 933 TGTATTATTTCTCTCTGGGCTCCCTACAAATGTCCTCTCTCTGACACACCTTCCAGGA 992
DB 250 eValTyrPheLeuPheThrProTyrAsnIleValIleLeuLeuAsnThrPheGlnG 270
QY 993 ATCTTTTGGCTGAAATGTCAGTGTCTTAACAGTGTGACAGCTGTGACAGCTATG 1052
DB 270 uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnVal 290
QY 1053 AGAGACTCTTGGGATGACGCTGCTGCATCAACCCATCATCTATGCTTGTCTGGGGA 1112
DB 290 rGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheValGlyG 310
QY 1113 GAAGTTCAGAACTACCTCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1172
DB 310 uLysPheArgTyrLeuSerValPhePheArgLysHisIleThrLysArgPheCysLy 330
QY 1173 ATGCTGTCTTATTTTCCAGCAAGAGGCTCCCGCAGCAGCAGCTCAGTTTACACCGGATC 1232
DB 330 sGlnCysProValPheTyrArgGluThrValAspGlyValThrSerThrAsnThrPro 350
QY 1233 CACTGGGAGCAGCAATATCTGTGGGCTG 1263
DB 350 rThrGlyGluGlnGluValSerAlaGlyLeu 360

RESULT 14

US-09-534-185-50
Sequence 50, Application US/09534185
Patent No. 6403767
GENERAL INFORMATION:
APPLICANT: Graham, Gerard J. et al.
TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
Heptahelical Receptor Superfamily and Uses
Therefor
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/534,185
FILING DATE: 24-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/045,583
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

| | | |
|-----|--|-----|
| 234 | ACAAGATGGATATCAAGTCTCAAGTCCAACTATGACATCAATATTATTATACATCCGAG | 293 |
| QY | : : | |
| 235 | ASnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla | 30 |
| Db | : : | |
| 236 | CCCTGCCAAAAATCAATGTGAAGCAAAATGCGAGCCGCTCTGCTCGCTCGCTCACTCA | 353 |
| QY | : : | |
| 237 | ProCysHisLysPheAspValysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer | 50 |
| Db | : : | |
| 238 | CTGGTGTTCATCTTTGGTGTTTGTGGCAACATCTGGTGCATCTCTCATCTCATAACTGC | 413 |
| QY | : : | |
| 239 | LeuValPheIlePheGlyPheValGlyAsnMetLeuValLeuLeuIleLeuAsnCys | 70 |
| Db | : : | |
| 240 | AAAAAGCTGAAGAGCATGACTGCATCTACTCTGCTCAACTGGCCATCTCTGACCTGTTT | 473 |
| QY | : : | |
| 241 | LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeu | 90 |
| Db | : : | |
| 242 | TTCCTCTTACTGTCCCTTCTGGCGTCTACTATGCTGCGCCAGTGGGACTTTGGAAAT | 533 |
| QY | : : | |
| 243 | PheLeuIleThrLeuProLeuTrpAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn | 110 |
| Db | : : | |
| 244 | ACAAATGTCTCACTCTTGACAGGGCTCTATTTTATAGCTTCTCTCTCGGAATCTCTTC | 593 |
| QY | : : | |
| 245 | AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhePhe | 130 |
| Db | : : | |
| 246 | ATCATCTTCCTGCACATCGATGAGTACTGTGGTGTCTGTCACATGCTGTGTTCCTTAAA | 653 |
| QY | : : | |
| 247 | IleIleLeuLeuThrIleAsnArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys | 150 |
| Db | : : | |
| 248 | GCCAGGACGGTCACTTTGGGGTGGTGACAAGTGTGATCACTTGGTGGTGGCTGTGTTT | 713 |
| QY | : : | |
| 249 | AlaArgThrValThrPheGlyValValThrSerValIleThrTrpLeuValAlaValPhe | 170 |
| Db | : : | |
| 250 | GGTCTCTCCAGGAATCACTTTTACAGATCTCAAAAAGAGGCTCTCATTCATCACCTGC | 773 |
| QY | : : | |
| 251 | AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys | 190 |
| Db | : : | |
| 252 | AGC-----TCTCATTTTCCATACATTAAAGATAGTACATCTGGG | 812 |
| QY | : : | |
| 253 | GlyProTyrPheProArgGlyTrpAsnAsn---PheHisThrIleMetArgAsnIleLeuGl | 210 |
| Db | : : | |
| 254 | GCTGTCTCTCGCGCTGCTTGTCATGGTCACTCTGCTACTCGGGAATCTTAAACTCTGCT | 872 |
| QY | : : | |
| 255 | IleuValLeuProIleuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeuLe | 230 |
| Db | : : | |
| 256 | TGGTGTCCGAATGAGAAGAGACGACGGCTGTGAGGCTTATCTCACCATCATGAT | 932 |
| QY | : : | |
| 257 | uArgCysArgAsnGluLysLysArgHisArgAlaValArgValIlePheThrIleMetIle | 250 |
| Db | : : | |

DB: 5

| | |
|------------------------|------------------|
| Alignment Scores: | |
| Pred. No.: | 9.96e-137 |
| Score: | 131.00 |
| Length: | 360 |
| Matches: | 258 |
| Percent Similarity: | 36 |
| Best Local Similarity: | Conservative: 36 |
| Query Match: | Mismatches: 46 |
| DB: | Indels: 11 |
| | Gaps: 3 |

US-09-938-703-3 (1-1442) x PCT-US95-00476-4 (1-360)

| | | | |
|----|------|--|------|
| QY | 234 | AAACAAGATGGATTATCAAGTCAGTCCAACTCATGACATCAATATTATATACATCGGAG | 293 |
| Db | 14 | AsnGluSerGlyGluGluValThrTrp----PhePheAspTyrAspTyr-----GlyAla | 30 |
| QY | 294 | CCCTGCGCAAAAAATCAATGTGAAGCAATCGACGCCGCTCCCTGCGCTCTACTCA | 353 |
| Db | 31 | ProCysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer | 50 |
| QY | 354 | CTGGTGTCATCTTTGGTTTTGTGGCAACATCGTGGTCATCCCTCATCTGATAAACTGC | 413 |
| Db | 51 | LeuValPheIlePheGlyPheValGlyAsnMetLeuValValLeuIleLeuIleAsnCys | 70 |
| QY | 414 | AAAAGGCTGAAGAGCATGACTGACATCTTACTCTGCTCAACCTGGCCATCTCTGACCTGTTT | 473 |
| Db | 71 | LysLysLeuLysCysLeuThrAspLeyLysLeuLeuAsnLeuAlaIleSerAspLeuLeu | 90 |
| QY | 474 | TTCTTCTTACTGTCOCCTCTGGGCTCAGTATGCTGCGCCGACAGTGGAGCTTTGGAAT | 533 |
| Db | 91 | PheLeuIleThrLeuProLeuTrpAlaHisSerAlaAlaAsnGluTrpValPheGlyAsn | 110 |
| QY | 534 | ACAAATGTGCAACTTCGACGGGCTCTATTTTATAGCTTCTCTCTGGAATCTTCTTC | 593 |
| Db | 111 | AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhePhe | 130 |
| QY | 594 | ATCATCCCTCTGCAATCCATAGTACTCGCTCGCTGCTCCATGCTGTGTTCTGTTTAAAA | 653 |
| Db | 131 | IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys | 150 |
| QY | 654 | GCACAGAGCGTCACCTTTGGGTGGTGACAAAGTGTGATCACTGGGTGGTGGCTGTGTTT | 713 |
| Db | 151 | AlaArgThrValThrPheGlyValValThrSerValIleThrTrpLeuValAlaValPhe | 170 |
| QY | 714 | GCCTCTCTCCACGAATCATCTTACAGATCTCAAAAGAGGCTCTTCATTCACACCTGC | 773 |
| Db | 171 | AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys | 190 |
| QY | 774 | AGC-----TCTCATTTTCCATCATTAAGATAGTCTATGCTTTGGG | 812 |
| Db | 191 | GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuGl | 210 |
| QY | 813 | GCATGTCCTGCGCGCTGTCATCGGTCACTCTCTACTCTCGGGAATCCTAAAACCTCGCT | 872 |
| Db | 210 | YleuValLeuProLeuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeuLe | 230 |
| QY | 873 | TCGGTGTGCAATGAGAAGAGACAGCGGTGTGAGGCTTATCTTCACCATCATGAT | 932 |
| Db | 230 | uArgCysArgAsnGluLysLysArgHisArgAlaValArgValIlePheThrIleMetI | 250 |
| QY | 933 | TGTTTATTTCTCTCTGGGCTCCCTACACATGTCTCTCTCTGTAACACCTTCCAGGA | 992 |
| Db | 250 | eValTyrPheLeuPheTrpThrProTyrAsnIleValIleLeuLeuAsnThrPheGlnGl | 270 |
| QY | 993 | ATTCTTTGGCTGAAATAATTCAGTGTCTTCAACAGGTGGACCAAGCTATGACAGTGAC | 1051 |
| Db | 270 | uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnValTh | 290 |
| QY | 1053 | AGAGACTCTGGGATGACGCAGCTGCTGCATCAACCCCATCATCTATGCTTTGTGCGGGA | 1111 |
| Db | 290 | rGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheValGlyGl | 310 |
| QY | 1113 | GAAGTTCAGAACTCACTCTTAGTCTTCTTCAAAGAGCATGTCGAACGCTTCTCGAA | 1171 |
| Db | 310 | uLysPheArgArgTyrLeuSerValPhePheArgLysHisIleThrLysArgPheCysLy | 330 |
| QY | 1173 | ATGCTCTTCTATTTTCCAGCAAGAGCGCTCCCGGACGCAAGCTCTATGTTTACACCCGATC | 1231 |
| Db | 330 | sgLncCysProValPheTyrArgGluThrValAspGlyValThrSerThrAsnThrProse | 350 |
| QY | 1233 | CATCGGGGAGCAGGAATATCTGTGGGCTTG | 1263 |
| Db | 350 | rThrGlyGluGlnGluValSerAlaGlyLeu | 360 |

Search completed: June 3, 2003, 18:52:47
Job time : 43.6947 secs

* 1743.3 00.4 03 03-013-033-13 sequence 13

| | | |
|--------------------------|---------------------------|-----------------|
| Alignment Scores: | | |
| Pred. No.: | 1.73e-158 | 352 |
| Score: | 1745.50 | 340 |
| Percent Similarity: | 96.60% | Matches: 1 |
| Best Local Similarity: | 96.33% | Conservative: 0 |
| Query Match: | 66.43% | Mismatches: 12 |
| DB: | 9 | Indels: 1 |
| | | Gaps: 1 |
| US-09-938-703-3 (1-1442) | x US-10-086-814-1 (1-352) | |


```

RESULT 2
US-09-734-221A-14
; Sequence 14, Application US/09734221A
; Publication No. US20030096221A1
; GENERAL INFORMATION:
APPLICANT: LITTMAN, DAN R.
; DENG, HONGKUI
; ELLMEIER, WILFRIED
; LANDAU, NATHANIEL R.
; LIU, RONG
;
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
; MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
; USES THEREOF
;
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/734,221A
; FILING DATE: 11-Dec-2000
; CLASSIFICATION: <Unknown>
;
; PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/666,020
; FILING DATE: 19-JUN-1996
; APPLICATION NUMBER: US 08/227,319
; FILING DATE: 13-APR-1994
;
; ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-004 N2
;
; TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
;
; INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
;
; ORIGINAL SOURCE:
ORGANISM: Homo sapiens
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-734-221A-14

Alignment Scores:
Pred. No.: 1,73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 9 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-734-221A-14 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATTATTATACATCGAGCCCTGC 299
;
;
;
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleasnTyrTyrThrSerGluProCys 20
;
;
;
QY 300 CAATAAATCAATGTGAAGCAAAATCGACGGCCCTCTCGCTCCGCTACTACTACTGGTG 359
;
;
;
Db 21 GlnGlysIleAsnValLysGlnIleAlaIaIardLeuLeuProLeuTyrSerIleuVal 40
;
;
;

```


Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIl 240
QY 927 CATGATGTTATTTCTCTCTGGCTCCCTACACATCTGCTCTCTCTCTGACACCTT 986
Db 240 eMetIleValTyPheLeuPheThrPalaProTyrrAsnIleValLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATTCAGTAGCTCTAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 280
QY 1047 GGTGACAGACTCTGGGATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPheVa 300
QY 1107 CGGGAGAGAGTTCAGAACTACCTCTTAGTCTCTTCCAAAGACATGCGCAAGCCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATAGCTGTTCTATTCTTCCAGCAAGAGGCTCCGAGCGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrrTh 340
QY 1227 CGGATCCACTGGGAGCAGGAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 4

US-09-813-653-15
; Sequence 15, Application US/09813653
; Patent No. US20020064770A1
; GENERAL INFORMATION:
; APPLICANT: Nestor, John
; APPLICANT: Wilson, Carol
; APPLICANT: See, Raymond
; APPLICANT: Tan Hehir, Christina
; TITLE OF INVENTION: Binding Compounds and Methods For Identifying Binding Compounds
; FILE REFERENCE: CMS-005
; CURRENT APPLICATION NUMBER: US/09/813,653
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,946
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/190,996
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,299
; PRIOR FILING DATE: 2000-03-21
; SOFTWARE: PatentIn version 3.0
; NUMBER OF SEQ ID NOS: 44
; SEQ ID NO 15
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-653-15

Alignment Scores:
Align. No.: 1-73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservatism: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-813-653-15 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATATATACATCGGAGCCCTGC 299
Db 1 MetaAspTyrrGlnValSerSerProIleTyrrAspIleAsnTyrrThrSerGluProCys 20
QY 300 CAARAATCAATGTGAGCAAAATCGAGCCGCCCTCTGCTGCTCTACTACTGCTGTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProProLeuTyrrSerLeuVal 40
QY 360 TTCATCTTGGTTTGTGGGCAACATCTGCTGCTCATCTCTCATCTCTGATATAAATGCAAAAGG 419

Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuAsnCysLysArg 60
QY 420 CTGAGACATGACTGACATCTACCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 61 LeuLysSerMetThrAspIleTyrrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTTACTGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
Db 81 LeuThrValProPheThrPalaHisTyrrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTTCAGAGGCTCTATTTATAGGCTCTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
QY 600 CTCTCGCAATCCGATAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrrLeuAlaValIleHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACTTTGGGCTGTGACAAAGTCTGATCCTTGGGTGGTGGTGGTGGTGGTGGT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTTCATTCATACCTGCACTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGATATGCTCAT 806
Db 181 HisPheProTyrrSerGlnTyrrGlnPheThrIlePheGlnThrLeu-LysIleValIl 200
QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrrSerGlyIleLeuLysTh 220
QY 867 TCTGCTCGGTGCTGCAAAATGAGAAAGAGGCTGAGGCTGAGGCTTATCTTCCCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrIl 240
QY 927 CATGATGTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrrPheLeuPheThrPalaProTyrrAsnIleValLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATAATTCAGTAGCTCTAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 280
QY 1047 GGTGACAGACTCTTGGGATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrrAlaPheVa 300
QY 1107 CGGGAGAGAGTTCAGAACTACCTCTTAGTCTCTTCCAAAGACATTCGCCAAAGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGTTCTATTCTTCCAGCAAGAGGCTCCGAGCGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrrTh 340
QY 1227 CGGATCCACTGGGAGCAGGAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 5

US-09-796-202-1
; Sequence 1, Application US/09796202
; Patent No. US20020068813A1
; GENERAL INFORMATION:
; APPLICANT: Dragic, Tatjana
; APPLICANT: Olson, William
; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION
; FILE REFERENCE: 2048/61010/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/796,202

; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: human
US-09-796-202-1

Alignment Scores:

Pred. No.: 1-73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservatives: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-796-202-1 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATATATACATCGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAAATCAATGTGAAGCAAAATCGCAGCCGCTCTGCTCCGCTCTACTCTACTGGTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGTGTCATCTCATCTCATCTGATAACTGCAAAAG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGAGCATGACATCTACCTGCTCAACCTGCGCATCTCTGACTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCTGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAAATG 539
Db 81 LeuThrValProPheIlePheAlaHisTyrAlaAlaGlnIlePhePheIleAsnMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTATPAGGCTTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGACATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspTyrLeuAlaValAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGTCACCTTTGGGTTGGTGGACAACTGTGATCACTGGGTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrIlePheValAlaValPheAlaSer 160
QY 720 CTCCTGAGTATCATCTTTACAGATCTCAAAAGAGGTCTTCATPACACCTGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTACAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeuLysIleValIle 200
QY 807 CTTGGGCTGGTCTCGCGCTGCTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 866
Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCGGTGTCAAAATGAGAGAGAGAGGACAGGCTGTGAGGCTTACTTCCACAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
QY 927 CATGATGTTTATTTCTCTGCTGGGCTCCCTACACATGCTCTTCTCTCCGACACCTT 986
Db 240 eMeIleValTyrPheLeuPheIlePheIlePheIlePheIlePheIlePheIlePheIle 260
QY 987 CCAGGAATCTTTGGCTGAATAATTGCAAGTACCTTCAACAGGTGGGACCAAGTATGCA 1046

Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 1166
QY 1047 GGTGACAGACATCTTTGGGATGACGACATCTGCTGATCAACCCCATCATCTATGCTTGT 1106
Db 280 nValThrGluThrLeuGlyMetIleHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTCTTCCAAAGCAATGCAAAAGCGTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGCTTATTTCCAGAGAGGCTCCCGAGCGAGCAGCTCAGTCTTACAC 1226
Db 320 eCysLysCysCysIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGGAATATCTGTGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluSerValGlyLeu 352

RESULT 6
US-09-938-719-5
; Sequence 5, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knodde, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA: US/09/938, 719
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-719-5

Alignment Scores:
Pred. No.: 1-73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservatives: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-938-719-5 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATATATACATCGGAGCCCTGC 299

Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAAATCATGATGAGCAATCGAGCGCCCTCTCGCTCGCTCTACTCTACTCTGGT 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGCAACATGCTGCTGCTCCTCATCTCTGATACCTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuLeuLeuLeuLeuLeu 60
QY 420 CTGAAGAGCATGATGATCATCTACTCTCAACCTGCGCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTCTTGGGCTCACTATGCTGCGCGCCAGTGGAGCTTTGGAATACATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGCAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGACATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAla 140
QY 660 ACGGTCACCTTTGGGCTGTCGACAGTGTGATCACTTGGTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCGAGGAATCATCTTACAGATCTCAACATCTCAAAAGAGCTCTCATCTACCTGCGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
QY 807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValIleProLeuLeuValMetValIleCysTyrSerGlyLeuLeuLysTh 220
QY 867 TCTGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATGCTTTATTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCCTGAATATGTCAGTACTTACAGGTTGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGACTCTGGGAGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAATGCTGCTTATTTTCCAGAGAGGCTCCGAGGAGGAGCTAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACATGGGACAGGAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 7

US-09-939-226-5

; Sequence 5, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMETIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-939-226-5

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 1,73e-158 | Length: | 352 |
| Score: | 1745.50 | Matches: | 340 |
| Percent Similarity: | 96.60% | Conservative: | 1 |
| Best Local Similarity: | 96.32% | Mismatches: | 0 |
| Query Match: | 66.42% | Indels: | 12 |
| DB: | 10 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x US-09-939-226-5 (1-352)

| | | | |
|----|-----|--|-----|
| QY | 240 | ATGATATTCAAGTGTCAAGTCCAATCTATGACATCAATATTATATACATCGAGCGCTGC | 299 |
| Db | 1 | MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys | 20 |
| QY | 300 | CAAAAAATCAATGTGAGCAAAATCGAGCGGCTCTCTGCTCGGCTCTACTCTACTGTGTG | 359 |
| Db | 21 | GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal | 40 |
| QY | 360 | TTCATCTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATAAATGCAAAAGG | 419 |
| Db | 41 | PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuLeuLeuLeuLeuLeu | 60 |
| QY | 420 | CTGAAGAGCATGATGATCATCTACTCTCAACCTGCGCATCTCTGACCTGTTTTCCTT | 479 |
| Db | 61 | LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu | 80 |
| QY | 480 | CTTACTGTCCTCTTGGGCTCACTATGCTGCGCGCCAGTGGAGCTTTGGAATACATG | 539 |
| Db | 81 | LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCAACTCTTGCAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC | 599 |

RESULT 8
US-09-938-703-5
; Sequence 5, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,703
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-938-703-5

Alignment Scores:
Pred. No.: 1,73e-158 Length: 352
Score: 1745.50 Matches: 340
Percent Similarity: 96.60% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 0
Query Match: 66.42% Indels: 12
DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-938-703-5 (1-352)
QY 240 ATGGATTCAAGTGCAGTCCCAATCTATGACATCAATATTATATACATCGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnIleTyrThrSerGlnProCys 20
QY 300 CAAAAATCAATGTGAAGCAAAATGCAGCCGCTCCTCGCTCTACTACTACTGGTG 359
Db 21 GLyLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCACTCTTGTTTGTGGCAACATSGTCACTCCCTCATCTGTAACTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleuIleAsnCysLysArg 60
QY 420 CTGAAGAGCATGACTGACATCTACTGCTCAACTGGGCATCTCGACCTGTTTCCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCTGGCTCAGTATGCTGCCGCCAGTGGGACATTGGAATACAAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaIleGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAACTGATAGGTACTGGCTGTCGTCATGCTGTGTTGCTTTAAAGCAGG 659
Db 121 LeuLeuThrIleLeuAspArgTyrLeuAlaValIleHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTCACTTTGGGGTGGTGACAAAGTGCATCAGTGGGTGGCTGTGTTGGCTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaPheAlaSer 160
QY 720 CTCCTCAGGAATCATCTTTACCAAGATCTCAAAAAGAGGTCTTCATTACACCTGCAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
QY 780 CATTTTCATAC-----ATTAAGCATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValIle 200
QY 807 CTTGGGGCTGCTCGCGCTGCTTGTCATGTCATCTGCTACTTCGGGAATCCATAAAAC 866

```

Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysYrSerGlyIleLeuLysnH 220
QY 867 TCTGCTTCGGTGTGCAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrI 240
QY 927 CATGATTGTTTATTTCTTCGCGGCTCCCPACAACTGTCTTCCTGCAACGCTT 986
Db 240 eMeIleValTyPheLeuPheThrAlaProTyAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGAAATCTTGGCTGATATTCGAGTACCTGCTACAGCTTGCAGCAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
QY 1047 GGTGACAGAGACTCTGGGATGACGACGCTGCTGCAACCCCATCATCTATGCTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPheVa 300
QY 1107 CGGGAGAACTTCAGAACTACCTCTTAGCTTCTTCCAAAGCAGCATTCGCAACGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAAAATGCTGTCTTATTTCCAGCAAGAGGCTCCGAGCAGCAAGCTCAGTTTAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyTh 340
QY 1227 CGCATTCACCTGGGAGCAGGAATATCTGTGGGCTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 9

US-10-106-623-2

; Sequence 2, Application US/10106623

; Patent No. US2002015088A1

; GENERAL INFORMATION:

; APPLICANT: Gray, Patrick W.

; Schweickart, Vicky L.

; Raport, Carol J.

; TITLE OF INVENTION: Chemokine Receptor Materials and Methods

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 S. Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/106,623

; FILING DATE: 26-Mar-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/771,276

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: No. US2002015088A1and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 27866/33670

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 352 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: /- "88C amino acid sequence"

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-106-623-2

Alignment Scores:

Pred. No.: 1,73e-158 Length: 352

Score: 1745.50 Matches: 340

Percent Similarity: 96.60% Conservative: 1

Best Local Similarity: 96.32% Mismatches: 0

Query Match: 66.42% Indels: 12

DB: 12 Gaps: 1

US-09-938-703-3 (1-1442) x US-10-106-623-2 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCATCTATGACATCAATATTATATCATCGGAGCCCTGC 299

Db 1 MetAspTyGlnValSerSerProIleTyAspIleAsnTyTyThrSerGluProCys 20

QY 300 CAAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCGCTCTACTCACTGGTG 359

Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuProLeuTySerLeuVal 40

QY 360 TTTCATCTTTGGTTTGGGCAACATGCTGGTCACTCTCATCTCTGATTAACATGCAAAAGG 419

Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60

QY 420 CTGAAGAGCATGCTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479

Db 61 LeuLysSerMetThrAspIleTyLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80

QY 480 CTCTACTGCCCTCTCTGGGCTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539

Db 81 LeuThrValProPheThrAlaHisTyAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100

QY 540 TGTCACCTCTTGACAGGCTCTATTTATAGGCTTCTTCTCTGGAATCTTCTTCATCATC 599

Db 101 CysGlnLeuLeuThrGlyLeuTyPheIleGlyPheSerGlyIlePhePheIleIle 120

QY 600 CTCTCGACATCATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659

Db 121 LeuLeuThrIleAspArgTyLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140

QY 660 AGGTCACCTTTGGGTTGGTGACAAAGTGTCATCTGGGTTGGGTTGGGTTGGGTTGGGTT 719

Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160

QY 720 CTCCCGAGCAATCATCTTACAGATCTCAAAAGAGGCTCTCATACACCTGCACTGCT 779

Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyThrCysSerSer 180

QY 780 CATTTTCCATAC-----ATTAAGATAGTATCAT 806

Db 181 HisPheProTySerGlnTyGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200

QY 807 CTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866

Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTySerGlyIleLeuLysTh 220

QY 867 TCTGCTTCGGTGTGCAAAATGAGAAGACAGGACAGGCTGTGAGGCTTATCTTCACCAT 926

Db 220 rLeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrI 240

QY 927 CATGATTCTTTATTTCTCTTCTGGGCTCCCTACAACTTCTCTCTCTCTCTCTCTCTCT 986

Db 240 eMeIleValTyPheLeuPheThrAlaProTyAsnIleValLeuLeuLeuAsnThrPh 260

QY 987 CCAGAAATCTTGGCTGATTAATTCAGTACTCTACAGGTTGGACCAAGCTATGCA 1046

Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280

QY 1047 GGTGACAGAGACTCTTGGGATGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106

| | | |
|------|--|------|
| 101 | CysGlnLeuThrGlyLeuTyPheIleGlyPheSerGlyIlePhePheIleIle | 120 |
| 600 | CTCCGTGCAATCATGATAGTACTGGCTGCTGCCATCTGCTGTTTAAAGCCAGG | 659 |
| 121 | LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg | 140 |
| 660 | ACGGTCACCTTTGGGTGGTGACAAAGTGATCACTTTGGGTGGTGCTGTTCGCTCT | 719 |
| 141 | ThrValThrPheGlyValValThrSerValIleThrTrpValAlaValPheAlaSer | 160 |
| 720 | CTCCGAGGAATCATCTTTACAGATCTCAAAAGAGGCTTCATTACACCTGCAGCTCT | 779 |
| 161 | LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyThrCysSerSer | 180 |
| 780 | CATTTTTCATAC-----ATAAAGATAGTGCAT | 806 |
| 181 | HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI | 200 |
| 807 | CTTGGGGTGGTCCGTCGGCTGTGCATGGTGCATCTGCTACTCGGAATCTCAAAAC | 866 |
| 200 | eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrseryIleLeuLysTh | 220 |
| 867 | CTGTCTTCGGTTCGAATAGAAAGAGAGGACAGAGGCTGAGGCTTATCTTCACAT | 926 |
| 220 | rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI | 240 |
| 927 | CATGATTGTTATTTTCTCTCTGGGTCCTCTACACATTTGCTCTCTCGTGAACCTT | 986 |
| 240 | eMetIleValTyPheLeuPheThrPalProTyrAsnIleValLeuLeuLeuAsnThrPh | 260 |
| 987 | CCAGGAATCTTTGGCGCTCAATAATTGCGTAGTCTTAACAGGTGGACCAAGCTATGCA | 1046 |
| 260 | eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG | 280 |
| 1047 | GETGACAGACACTTTGGGATGAGCAGCTGCTGCATCAACCCCATCTATGCGCTTGT | 1106 |
| 280 | nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyralaPheVa | 300 |
| 1107 | CGGGAGAGATTTCAGAACTACCTTCCTTAGTCTCTTCTCCAAAGCACATTTGCCAAGCTT | 1166 |
| 300 | IleGlyIlyPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgph | 320 |
| 1167 | CTGCAATGCTGTTCATTTTCCAGCAAGAGGCTCCGCGAGGAGCAAGCTCAGTTTACAC | 1226 |
| 320 | eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyTh | 340 |
| 1227 | CGATCCACTGGGACGACGAAATATCTGTGGGCTTG | 1263 |
| 340 | rArgSerThrGlyGlnGlnGluIleSerValGlyLeu | 352 |

RESULT 11
US-10-232-686-2
; Sequence 2, Application US/10232686
; Publication No. US20030023044A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human G-Protein Chemokine Receptor (CCR5) HDGKR10
; FILE REFERENCE: 1488-115000N
; CURRENT APPLICATION NUMBER: US/10/232,686
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-232-686-2

Alignment Scores:

Pred. No.: 1,01e-157 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservativity: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 9 Gaps: 1

US-09-938-703-3 (1-1442) x US-10-232-686-2 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAATCAATGTGAGCAATCGCAGCCGCTCGCTCGCTCTACTCATCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCACTTTGGTTTGTGGCAACATGCTGCTCATCTCATCTGATAAATGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAGACATGACTGACATCTACTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTCTGGCTCCTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
Db 81 LeuThrValProPheThrPheAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTGTGACGGCTCTATTATAGGCTTCTCTGGAATCTTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleCysPheSerGlyIlePhePheIleIle 120
QY 600 CTCTGACATCATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValAlaHisAlaValPheAlaLeuLysAlaArg 140
QY 660 AGGTCACCTTTGGGTGGTGACAGTGTGATCCTGCTGCTGCTGCTGCTGCTGCTG 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGCAATCATCTTACAGATCTCAAAAGAGGTCTCTATACCTGCGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheThrPheLysAsnPheGlnThrLeu-LysIleValI 200
QY 807 CTGCGGCTGGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTCGGTGTCGAATGAGAGAGACAGCGGCTGTGAGGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrIl 240
QY 927 CATGATTCTTTATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
Db 240 eMetIleValTyrPheLeuPheThrPheAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTGAATTAATGCTAGTCTTAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG1 280
QY 1047 GTGACAGACTCTTGGAGTACGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
```

```
QY 1107 CGGGAGAGTTCAGAACTACTCTTAGTCTTCTCCAAAAGACATTCGCAACGCTT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTCAAAATGCTGTCTATTTCACGAAAGCTCCGAGCGGCAAGCTCAGTTTACAC 1226
Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGCTG 1263
Db 340 rArgSerThrGluGluGlnGlnIleSerValGlyLeu 352
```

RESULT 12

US-09-725-285-2
Sequence 2, Application US/09725285
Patent No. US20010000241A1
GENERAL INFORMATION:
APPLICANT: Li, Yi
TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGNR10
FILE REFERENCE: 1488, 1150003
CURRENT APPLICATION NUMBER: US/09/725,285
CURRENT FILING DATE: 2000-11-29
PRIOR APPLICATION NUMBER: 09/339,912
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/195,662
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 08/466,343
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
TYPE: PRT
LENGTH: 352
ORGANISM: Artificial Sequence: Genomic
FEATURE:
OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-725-285-2

Alignment Scores:

Pred. No.: 1,01e-157 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservativity: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-725-285-2 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAATCAATGTGAGCAATCGCAGCCGCTCGCTCGCTCTACTCATCTGCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCACTTTGGTTTGTGGCAACATGCTGCTCATCTCATCTGATAAATGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAGACATGACTGACATCTACTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTCTGGCTCCTACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
Db 81 LeuThrValProPheThrPheAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGGAGTACGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleCysPheSerGlyIlePhePheIleIle 120
```

```

600 CTCTGACAACTAGTAGTACCTGGCTGCTGCTCCATGCTGTGTGTTGCTTTAAAGCAGG 659
121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
560 ACGGTACCTTTGGGTGGTGACAGTGTGATCAGTGGTGGTGGTGGTGGTGGTGGTGGT 719
141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
720 CTCCAGGAATCACTTTACAGATCTCAAAAGAGAGGCTTCATACACCTCCAGCTCT 779
161 LeuProGlyIleLeuPheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
780 CATTTCATATC-----ATTAAAGATAGTCAAT 806
181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleVal 200
807 CTGGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 866
200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
867 TCTGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 926
220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThr 240
927 CATGATGTTTATTTCTCTCTGGGCTCCCTACACATGTCCTCTCTCTCTCTCTCTCT 986
240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThr 260
987 CCAGGAATCTTTGGCTGAATAATGTCAGTGTCTTAACAGTTGGACCAAGCTATGCA 1046
260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMet 280
1047 GGTGACAGACTCTTGGGATGACACTGCTGATCAACCCATCATCTATGCTTGTCT 1106
280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPhe 300
1107 CGGGAGAGAGTTCAGAACTACTCTTGTAGTCTTCTTCCAAAGACATTCGCAAGCTT 1166
300 lGlyLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArg 320
1167 CTCAAATGCTGTCTATTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTACAC 1226
320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrThr 340
1227 CCAGTCCACTGGGAGCAGGAATATCTGGGGCTTG 1263
340 rArgSerThrGluGluGlnIleSerValGlyLeu 352

RESULT 13
US-09-779-879A-22
; Sequence 22, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Roschke, Viktor
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven, M.
; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
; FILE REFERENCE: 1488.115000A
; CURRENT APPLICATION NUMBER: US/09/779,879A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,258
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US 60/187,999
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 60/234,336
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 22
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-879A-22
Alignment Scores: 1,01e-157 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 10 Gaps: 1
US-09-938-703-3 (1-1442) x US-09-779-879A-22 (1-352)
QY 240 ATGATTATCAAGTGTCAAGTCCAACTATGATCACTAATATTATATACATCGAGCCCTGC 299
DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAAAATGCGAGCCGCTCCCTGCTCCGCTCTACTACTCGGTG 359
DB 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGTTTGTGGCAACATGCTGGTCATCTCATCTCTCATCTGATTAACCTGCA 419
DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuLeuLeuLeuLeuLeu 60
QY 420 CTGAAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTTCTGGGCTCAGTATGTCGGCCAGTGGGACTTGGAAATACATG 539
DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTTATAGCTTCTTCTGGAATCTTCTCATCATC 599
DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGCAATCAATAGTAGTACTGCTGCTCCATGCTGTGTTTTCCTTTAAAGCAGG 659
DB 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTACCTTTGGGTGGTGACAGTGTGATCAGTGGTGGTGGTGGTGGTGGTGGTGGT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCACTTTTACAGATCTCAAAAGAGAGGCTTTCATACACCTCCAGCTCT 779
DB 161 LeuProGlyIleLeuPheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
QY 780 CATTTCATATC-----ATTAAAGATAGTCAAT 806
DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleVal 200
QY 807 CTGGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 866
DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
QY 867 TCTGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 926
DB 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThr 240
QY 927 CATGATGTTTATTTCTCTCTGGGCTCCCTACACATGTCCTCTCTCTCTCTCTCTCT 986
DB 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThr 260
QY 987 CCAGGAATCTTTGGCTGAATAATGTCAGTGTCTTAACAGTTGGACCAAGCTATGCA 1046
DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMet 280
1047 GGTGACAGACTCTTGGGATGACACTGCTGATCAACCCATCATCTATGCTTGTCT 1106
DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPhe 300
1107 CGGGAGAGAGTTCAGAACTACTCTTGTAGTCTTCTTCCAAAGACATTCGCAAGCTT 1166
DB 300 lGlyLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArg 320
1167 CTCAAATGCTGTCTATTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTACAC 1226
DB 320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerValTyrThr 340
1227 CCAGTCCACTGGGAGCAGGAATATCTGGGGCTTG 1263
DB 340 rArgSerThrGluGluGlnIleSerValGlyLeu 352

; ORGANISM: Homo sapiens
US-09-779-879A-22
Alignment Scores: 1,01e-157 Length: 352
Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 10 Gaps: 1
US-09-938-703-3 (1-1442) x US-09-779-879A-22 (1-352)
QY 240 ATGATTATCAAGTGTCAAGTCCAACTATGATCACTAATATTATATACATCGAGCCCTGC 299
DB 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAAAATGCGAGCCGCTCCCTGCTCCGCTCTACTACTCGGTG 359
DB 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGTTTGTGGCAACATGCTGGTCATCTCATCTCTCATCTGATTAACCTGCA 419
DB 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuLeuLeuLeuLeuLeuLeu 60
QY 420 CTGAAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
DB 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTTCTGGGCTCAGTATGTCGGCCAGTGGGACTTGGAAATACATG 539
DB 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTTATAGCTTCTTCTGGAATCTTCTCATCATC 599
DB 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
QY 600 CTCCTGCAATCAATAGTAGTACTGCTGCTCCATGCTGTGTTTTCCTTTAAAGCAGG 659
DB 121 LeuLeuThrIleAspArgTyrLeuAlaValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGGTACCTTTGGGTGGTGACAGTGTGATCAGTGGTGGTGGTGGTGGTGGTGGTGGT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCACTTTTACAGATCTCAAAAGAGAGGCTTTCATACACCTCCAGCTCT 779
DB 161 LeuProGlyIleLeuPheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
QY 780 CATTTCATATC-----ATTAAAGATAGTCAAT 806
DB 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleVal 200
QY 807 CTGGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 866
DB 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLys 220
QY 867 TCTGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 926
DB 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuIlePheThr 240
QY 927 CATGATGTTTATTTCTCTCTGGGCTCCCTACACATGTCCTCTCTCTCTCTCTCTCT 986
DB 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThr 260
QY 987 CCAGGAATCTTTGGCTGAATAATGTCAGTGTCTTAACAGTTGGACCAAGCTATGCA 1046
DB 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMet 280
1047 GGTGACAGACTCTTGGGATGACACTGCTGATCAACCCATCATCTATGCTTGTCT 1106
DB 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPhe 300
```

QY 1107 CGGGAGAGTTTCAAGAACTACCTTAGTCTTCTTCCAAAAGACACATTGCCAAAGCGTT 1166
 Db 300 LGlyGlyLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAAACTGCTTCTATTTCCAGACAGAGGCTCCGGAGCAGAGCACTCAGTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnLysAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGluGlnGluLysSerValGlyLeu 352
 RESULT 14
 US-09-779-880A-22
 ; Sequence 22, Application US/09779880A
 ; Patent No. US20020061834A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, Craig A.
 ; APPLICANT: Roschke, Viktor
 ; APPLICANT: Li, Yi
 ; APPLICANT: Ruben, Steven, M.
 ; TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
 ; FILE REFERENCE: 1488.115000C
 ; CURRENT APPLICATION NUMBER: US/09/779,880A
 ; CURRENT FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: US 60/181,258
 ; PRIOR FILING DATE: 2000-02-09
 ; PRIOR APPLICATION NUMBER: US 60/187,999
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: US 60/234,336
 ; PRIOR FILING DATE: 2000-09-22
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-779-880A-22
 Alignment Scores:
 Pred. No.: 1.01e-157 Length: 352
 Score: 1737.50 Matches: 339
 Percent Similarity: 96.32% Conservative: 1
 Best Local Similarity: 96.03% Mismatches: 1
 Query Match: 66.11% Indels: 12
 DB: 10 Gaps: 1
 US-09-938-703-3 (1-1442) x US-09-779-880A-22 (1-352)
 QY 240 ATGGATATCAAGTCTCAAGTCCCAATCTATGACATCAATATTATATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
 QY 300 CAAAATCAATGTCAAGCAAAATCGAGCCGCTCTGCTCGCTCGCTCTACTCACTGGTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTGATCTTTGGTTTGGGCAACATCGTGTGTCATCTCATCTCATCTGATAAATCGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACATGACATCTACCTGCTCACTGCGCCATCTCTGACCTCTTTTCCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTCTGGGCTCACTATGTCGCCCGCCAGCTGGGACTTTGGAATACAAATG 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCACTCTTACAGAGGCTCTATTTATAGGCTTCTCTCTGGAAATCTTCTTCATCAGC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120

QY 600 CTCCTGCAAAATCGATAGTGTACCTGGCTGTGCTCCATGCTGTGTGTTGCTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrIleAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGCGTCACTTTGGGTGGTGTGACAAAGTGTGATCACTTGGGTGGTGGTGTGTTGCGTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCGCCAGAAATCATCTTTACCATGATCTCAAAAAGAGGTCTTCACTACCTGACGTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSer 180
 QY 780 CATTTTCCATAT-----ATTAAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
 QY 807 CTGGGGCTGCTCCGCGCTGCTGTGTGTCATGTCTACTCGGGAATCCTTAAAGAC 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCGCTGTGCAAAATGAGAAAGAGGACAGGCTGTGAGGCTTATCTTCACCAT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATTGTTTATTTTCTTCTGCGCTCCCTACACCATTCCTCTCTCTCTCTCTCTCT 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCCCTGGAATATTCAGTAGCTTACAGGTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGAGACTTCTGGGATGACGACTCTGTCATCAACCCATCATCTATGCTTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGTTCAGAACTACCTCTTGTAGTCTTCTTCCAAAAGCACATTCGCAACGCTT 1166
 Db 300 LGlyGlyLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAAAATGCTGTCTATTTTCCAGCAAGAGGCTCCGAGCAGGAGCTCAGTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnLysAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGluGlnGluLysSerValGlyLeu 352
 RESULT 15
 US-09-195-662A-2
 ; Sequence 2, Application US/09195662A
 ; Patent No. US20020076745A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Yi
 ; APPLICANT: Ruben, Steven, M.
 ; TITLE OF INVENTION: Human G-Protein Chemokine Receptor HDGNR10 (CCR5 Receptor)
 ; FILE REFERENCE: 1488.1150002
 ; CURRENT APPLICATION NUMBER: US/09/195,662A
 ; CURRENT FILING DATE: 1998-11-18
 ; PRIOR APPLICATION NUMBER: 08/466,343
 ; PRIOR FILING DATE: 1995-06-06
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence: Genomic
 ; FEATURE:
 ; OTHER INFORMATION: Deduced Amino Acid Sequence
 US-09-195-662A-2
 Alignment Scores:
 Pred. No.: 1.01e-157 Length: 352

Score: 1737.50 Matches: 339
Percent Similarity: 96.32% Conservative: 1
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 66.11% Indels: 12
DB: 10 Gaps: 1

US-09-938-703-3 (1-1442) x US-09-195-662A-2 (1-352)

```
QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProIleTyrAspIleAsnTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGACAGCAAAATCGAGCCCGCTCTCTCCCTCTACTCTACTCTGCTG 359
Db 21 GlnLysIleAsnValysGlnIleAlaAlaArgLeuLeuProIleuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGGGCAACATGCTGTGGTCTCATCTCTCATCTGATTAACCTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysLysArg 60
QY 420 CTGAGAGCATGACATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPhePheLeu 80
QY 480 CTACTGTCCCTCTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGSCCTATTTTATPAGSCTTCTCTGTGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATGATAGTAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 ACGCTCACCTTTGGGCTGGTGCACAGTGTGATCATCTGGTGGTGGTGGTGGTGGTGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCGCCAGAAATCATTTTACCAGATCTCAAAAGAAGTCTTCAATACACCTGCAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle 200
QY 807 CTGGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysThr 220
QY 867 TCTGCTTCGGTCTCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
QY 927 CATGATGTTTATTTCTCTCTGGGCTCCCTTACACATGCTCTCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCCTCAATATTGCAAGTAGCTCTACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSerSer 280
QY 1047 GGTGACAGAGACTCTTGGGATGAGCAGCTGCTGATCATCAACCCCATCATCTATGCTTTGT 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAGAGTTCAGAACTACCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1166
Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
QY 1167 CTGCAAAATGCTGTTCTATTTTCCAGCAAGAGGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
```

```
Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
QY 1227 CGATCCACTGGGGAGCAGCAAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGluGluGlnGlnIleSerValGlyLeu 352
```

Search completed: June 3, 2003, 19:13:30
Job time : 83.7773 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OK nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:42:09 ; Search time 50.6953 seconds
(without alignments)
5468.980 Million cell updates/sec

Title: US-09-938-703-3

Perfect score: 2828

Sequence: 1 GAATTCGCCCAACAGACCA.....AGTAGATTAGATCCGAATTC 1442

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-Q=/cg21/USPTO_spo1/US9938703/runat_03062003_161407_22855/app.query.fasta_1.1870
-DB=PIR73 -QFMT=Eastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USUR=US9938703 -CGN_1_1_89 -runat_03062003_161407_22855 -NCPU=6
-NO_MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: PIR73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|---------|--------------------|
| 1 | 1745.5 | 66.4 | 352 | A43113 | chemokine (C-C) re |
| 2 | 1331 | 50.6 | 360 | JC2443 | chemokine (C-C) re |
| 3 | 1191 | 45.3 | 374 | 138450 | chemokine (C-C) re |
| 4 | 1003.5 | 38.2 | 355 | A45177 | chemokine (C-C) re |
| 5 | 960 | 36.5 | 359 | 149341 | MIP-1 alpha recept |
| 6 | 953 | 36.3 | 355 | 2149339 | macrophage inflam |
| 7 | 909 | 34.6 | 355 | 2G02436 | chemokine (C-C) re |
| 8 | 891.5 | 33.9 | 360 | 2A57160 | chemokine (C-C) re |
| 9 | 878.5 | 33.4 | 360 | 2JC4587 | chemokine (C-C) re |
| 10 | 797.5 | 30.3 | 383 | 2S5594 | G protein-coupled |
| 11 | 776 | 29.5 | 356 | 2149340 | MIP-1 alpha recept |
| 12 | 716 | 27.2 | 355 | 2JC5067 | G protein-coupled |
| 13 | 688.5 | 26.2 | 354 | 2I58186 | probable G protein |
| 14 | 644 | 24.5 | 355 | 2JC4304 | orphan G protein-c |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 15 | 638.5 | 24.3 | 344 | 2 | JC5942 | chemokine receptor |
| 16 | 573.5 | 21.8 | 378 | 2 | B55735 | lymphocyte-specifi |
| 17 | 559.5 | 21.3 | 378 | 2 | A45680 | G protein-coupled |
| 18 | 556.5 | 21.2 | 378 | 2 | A55735 | G protein-coupled |
| 19 | 552.5 | 21.0 | 369 | 2 | JC5068 | G protein-coupled |
| 20 | 528.5 | 20.1 | 333 | 2 | 165989 | interleukin-8 rece |
| 21 | 511.5 | 19.5 | 360 | 2 | A53611 | interleukin-8 rece |
| 22 | 507 | 19.3 | 359 | 2 | A48921 | interleukin-8 rece |
| 23 | 505 | 19.2 | 355 | 2 | JQ1231 | interleukin-8 rece |
| 24 | 501.5 | 19.1 | 367 | 2 | JE0349 | interferon-inducib |
| 25 | 501 | 18.1 | 358 | 2 | A53752 | interleukin-8 rece |
| 26 | 498 | 18.9 | 350 | 2 | JN0621 | G protein-coupled |
| 27 | 492.5 | 18.7 | 350 | 2 | A39445 | interleukin-8 rece |
| 28 | 491.5 | 18.7 | 352 | 2 | G00048 | fusin (LESTRA) - c |
| 29 | 491.5 | 18.7 | 352 | 2 | A45747 | neuropeptide Y pep |
| 30 | 486 | 18.5 | 356 | 2 | S42096 | interleukin-8 rece |
| 31 | 483.5 | 18.4 | 353 | 2 | S28787 | neuropeptide Y pep |
| 32 | 468 | 17.8 | 374 | 2 | S42628 | G protein-coupled |
| 33 | 463 | 17.6 | 359 | 2 | A42656 | angiotensin II rec |
| 34 | 462 | 17.6 | 359 | 2 | JC2134 | angiotensin II rec |
| 35 | 460 | 17.5 | 359 | 2 | S44425 | angiotensin II rec |
| 36 | 459 | 17.5 | 359 | 2 | S15403 | angiotensin II rec |
| 37 | 458 | 17.4 | 359 | 2 | I51372 | angiotensin II rec |
| 38 | 457 | 17.4 | 359 | 2 | JH0621 | angiotensin II rec |
| 39 | 455.5 | 17.3 | 327 | 2 | S56162 | MDCR15 protein - h |
| 40 | 455.5 | 17.3 | 372 | 2 | S26667 | G protein-coupled |
| 41 | 454 | 17.3 | 359 | 2 | JQ1516 | angiotensin II rec |
| 42 | 453.5 | 17.3 | 374 | 2 | S32785 | G protein-coupled |
| 43 | 453 | 17.2 | 359 | 2 | JC1104 | angiotensin II rec |
| 44 | 452 | 17.2 | 359 | 2 | JC1194 | angiotensin II rec |
| 45 | 452 | 17.2 | 359 | 2 | A48857 | angiotensin II rec |

ALIGNMENTS

RESULT 1

A43113

chemokine (C-C) receptor 5 - human

N.Alternate names: C-C CKR-5; CCR5

C.Species: Homo sapiens (man)

C.Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text change 20-Jun-2000

C.Accession: A43113; S71808; A58834; A58832; G02653; A58833

R.Sanson, M.; Labbe, O.; Mollereau, C.; Vassart, M.

A.Title: Molecular cloning and functional expression of a new human CC-chemokine rec

A.Reference number: A43113; MUID:96241590; PMID:8639485

A.Accession: A43113

A.Molecule type: mRNA

A.Residues: 1-352 <SAM1>

A.Cross-references: GB:X91492; NID:g1262810; PIDN:CAA62796.1; PID:g1262811

R.Sanson, M.; Libert, F.; Doranz, B.J.; Rucker, J.; Liesnard, C.; Farber, C.M.; Sazac

M.; Imai, T.; Rana, S.; Yi, Y.; Smyth, R.J.; Collman, R.G.; Doms, R.W.; Vassart, G.;

Nature 382, 722-725, 1996

A.Title: Resistance to HIV-1 infection in caucasian individuals bearing mutant allele

A.Reference number: S71808; MUID:96345670; PMID:8751444

A.Accession: S71808

A.Status: nucleic acid sequence not shown; not compared with conceptual translation

A.Molecule type: DNA

A.Residues: 182-206; 207-230 <SAM2>

A.Accession: A58834

A.Status: nucleic acid sequence not shown; not compared with conceptual translation

A.Molecule type: DNA

A.Residues: 1-184; 'IKDSHIGAGPAACHGHLGNPKNSASVK' <SAM3>

A.Cross-references: GB:X99393; NID:g1524062; PIDN:CAA67767.1; PID:g1524063

A.Note: this frameshift mutation results in a non-functional receptor but confers a c

nd may have had a selective advantage by conferring resistance to Yersinia plague inf

R.Combadiere, C.; Ahuja, S.K.; Tiffany, H.L.; Murphy, P.M.

J. Leukoc. Biol. 60, 147-152, 1996

A.Title: Cloning and functional expression of CC CKR5, a human monocyte CC chemokine

A.Reference number: A58832; MUID:96295970; PMID:8699119

A.Accession: A58832

A.Molecule type: mRNA

A.Residues: 1-352 <COM1>

A:Cross-references: GB:U57840; NID:q1502408; PIDN:AABI7071.1; PID:q1502409
 A:Experimental source: clone 8, endotoxin-stimulated peripheral blood monocytes
 R:Combadriere, C.
 submitted to the EMBL Data Library, May 1996
 A:Reference number: H01541
 A:Accession: G02653
 A:Status: translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-89, 'L', 91-352 <COM2>
 A:Cross-references: EMBL:U57840
 J:Report, C.J.; Gosling, J.; Schweickart, V.L.; Gray, P.W.; Charo, I.F.
 J. Biol. Chem. 271, 17161-17166, 1996
 A:Title: Molecular cloning and functional characterization of a novel human CC chemokine
 A:Reference number: A58833; MUID:96291862; PMID:8663314
 A:Accession: A58833
 A:Molecule type: mRNA
 A:Residues: 1-352 <RAP>
 A:Cross-references: GB:U54994; NID:q1457945; PIDN:AKC50598.1; PID:q1457946
 C:Comment: This is a receptor for chemokines MIP-1alpha (see PIR:A30574), MIP-1beta (see
 C:Comment: Macrophage- and dual-tropic strains of HIV-1 bind to a complex of chemokine
 C:Genetics:
 A:Gene: GDB:CMK3R5; CCR5; CCR-5; CC-CR-5; CCR5; ChemR13
 A:Cross-references: GDB:I230510; OMIM:601373
 A:Map position: 3p21-3p21
 C:Function:
 A:Description: G protein-coupled receptor for chemokines MIP-1alpha, MIP-1beta and RANTES
 A:Note: Probably acts to control granulocyte proliferation and differentiation
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: AIDS; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembran
 F:32-56/Domain: transmembrane #status predicted <TM1>
 F:67-87/Domain: transmembrane #status predicted <TM2>
 F:103-124/Domain: transmembrane #status predicted <TM3>
 F:143-166/Domain: transmembrane #status predicted <TM4>
 F:193-218/Domain: transmembrane #status predicted <TM5>
 F:236-257/Domain: transmembrane #status predicted <TM6>
 F:285-300/Domain: transmembrane #status predicted <TM7>
 F:20-269,101-178/disulfide bonds: #status predicted
 F:268/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:336,337,342/Binding site: phosphate (Ser) (covalent) #status predicted
 F:340,343/Binding site: phosphate (Thr) (covalent) #status predicted

Alignment Scores:

| Pred. No.: | 6,41e-146 | Length: | 352 |
|------------------------|-----------|---------------|-----|
| Score: | 1745.50 | Matches: | 340 |
| Percent Similarity: | 96.60% | Conservative: | 1 |
| Best Local Similarity: | 96.32% | Mismatches: | 0 |
| Query Match: | 66.42% | Indels: | 12 |
| DB: | 2 | Gaps: | 1 |

US-09-938-703-3 (1-1442) x A43113 (1-352)

QY 240 ATGGATATCAAGTCAAGTCCAACTATGACATCAATATATATACATCGAGCGCTGC 299
 Db 1 MetAspTyrGlnValSerProIleTyrAspIleAsnTyrThrSerGluProCys 20
 QY 300 CAAATAATCAATGAAGCAATCGCAGCGCCCTCGCTGCTACTCTACTCTGCTGGT 359
 Db 21 GlnYsIleAsnValIysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTGGTTTGGGCAACATGCTGGTCATCTCTCATCTCTGATAACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysIysArg 60
 QY 420 CTGAAGACATGACATGACATCTACCTGCTCACCTGGCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuIysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTCTCCCTTCTGGCTCTCACTATGCTGCGCCAGTGGAGCTTTGGAAATACATG 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaGlnIleAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGCGCTCAATTTATAGGCTTCTCTCTGGAATCTTCTTCATCATC 599
 |||||||

Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGACATCGATAGTACCTGGCTGCTGCTCATCTGCTGTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGTCACTCTGGGCTGGTGCACAGTGTGATCACTTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCGAGGAATCATCTTTACAGATCTCAAAAGAGGCTTTCATTACACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTATCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValI 200
 QY 807 CTTGGGCTGGTCTCTGCTGGCTGCTGCTCATCTGCTACTCGGGAATCTTAAAC 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTCGGTGTCGAATGAGAGGACGACAGGCTGTGAGGCTTATCTTACCAT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATGATTGTTTATTTCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 280
 QY 1047 GGTGACAGACTCTTGGGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGGAGAGTTCAGAACTACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
 Db 300 lGlyGlnLysPheArgAsnTyrLeuValPheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTCGAAATGCTGTTCTATTTTCCAGAGCTCCCGAGGCTCCGAGGCTGCTGCTGCTGCTGCT 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGGAGGAATATCTGCTGGCTTG 1263
 Db 340 rArgSerThrGlyGlnGlnIleSerValGlyLeu 352

RESULT 2
 JC2443
 chemokine (C-C) receptor 2, splice form B - human
 N:Alternate names: C-C CR-2; monocyte chemoattractant protein 1 receptor; monocyte c
 C:Species: Homo sapiens (man)
 C>Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 20-Jun-2000
 C:Accession: JC2443; I38463
 R:Yamagami, S.; Tokuda, Y.; Ishii, K.; Tanaka, H.; Endo, N.
 Biochem. Biophys. Res. Commun. 202, 1156-1162, 1994
 A:Title: cDNA cloning and functional expression of a human monocyte chemoattractant
 A:Reference number: A53477; MUID:94195821; PMID:8146186
 A:Accession: I38463
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-360 <RES>

A;Cross-references: EMBL:U03905; NID:g472557; PIDN:AAA19120.1; PID:g472558

C;Genetics:

A;Gene: GDB:CMKBR2

A;Cross-references: GDB:337364; OMIM:601267

A;Map position: 3p21-3p21

C;Superfamily: vertebrate rhodopsin

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembran

F;43-70/Domain: transmembrane #status predicted <TM1>

F;81-100/Domain: transmembrane #status predicted <TM2>

F;115-136/Domain: transmembrane #status predicted <TM3>

F;154-178/Domain: transmembrane #status predicted <TM4>

F;207-226/Domain: transmembrane #status predicted <TM5>

F;244-268/Domain: transmembrane #status predicted <TM6>

F;287-309/Domain: transmembrane #status predicted <TM7>

F;14/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;113-190/Disulfide bonds: #status predicted

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 2,61e-109 | Length: | 360 |
| Score: | 1331.00 | Matches: | 258 |
| Percent Similarity: | 83.76% | Conservative: | 36 |
| Best Local Similarity: | 73.50% | Mismatches: | 46 |
| Query Match: | 50.65% | Indels: | 11 |
| DB: | 2 | Gaps: | 3 |

US-09-938-703-3 (1-1442) x JC2443 (1-360)

| | | | |
|----|-----|--|-----|
| QY | 234 | ACAAGATGATATCAAGTCCAGTCCCAATCATCATCAATATATTATACATCGGAG | 293 |
| DB | 14 | AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla | 30 |
| QY | 294 | CCCTGCCAAATATCATGTGACAGCAATTCGACCCCGCTCTCGCTCGCTCTACTCA | 353 |
| DB | 31 | ProCysHisLysPheAspValLysGlnIleGlyAlaGlnLeuLeuProLeuTyrSer | 50 |
| QY | 354 | CTGGTCTTCATCTTGGTGTGGGCAACATGCTGCTCATCTCATCTCATATAAATCG | 413 |
| DB | 51 | LeuValPheIlePheGlyPheValGlyAsnMetLeuValValleuLeuIleAsnCys | 70 |
| QY | 414 | AAAGCTGACAGACATGACTGACATCTACCTGCTCAACCTGCCATCTCTGACCTGTT | 473 |
| DB | 71 | LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeu | 90 |
| QY | 474 | TTCTCTTCTACTGCTCCCTCTCGGCTCACTATGCTCGCCGCTGGAGTTTGAAT | 533 |
| DB | 91 | PheLeuIleThrLeuProLeuThrAlaHisSerAlaAsnGluThrValPheGlyAsn | 110 |
| QY | 534 | ACAATGTCTCACTCTGACAGGCTCTATTATTAGGCTTCTTCTCTGGAATCTCTTC | 593 |
| DB | 111 | AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhePhe | 130 |
| QY | 594 | ATCATCTCTGACAAATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA | 653 |
| DB | 131 | IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys | 150 |
| QY | 654 | GCCAGGACGCTCACCTTGGGTGGTGAAGTGTGATGATGATGATGATGATGATGAT | 713 |
| DB | 151 | AlaArgThrValThrPheGlyValValThrSerValIleThrIlePheValAlaValPhe | 170 |
| QY | 714 | GCCTCTCTCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTTCTCATCATCATCTG | 773 |
| DB | 171 | AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys | 190 |
| QY | 774 | AGC-----TCTCATTTTCCATACATTAAGATAGATAGATAGATAGATAGATAG | 812 |
| DB | 191 | GlyProTyrPheProArgGlyTyrAsnAsn-PheHisThrIleMetArgAsnIleLeuG | 210 |
| QY | 813 | GCTGCTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 872 |
| DB | 210 | YleuValleuProLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeuLeu | 230 |
| QY | 873 | TCGGTCTCGAATGAGAGAGGACAGGCTGTGAGGCTTATCTTTCACCATCATGAT | 932 |

| | | | |
|----|------|--|------|
| DB | 230 | uArgCysArgAsnGluLysArgHisArgAlaValArgValIlePheThrIleMetI | 250 |
| QY | 933 | TGTTTATTCTCTCTGGGCTCCCTACACATGCTCTCTCTCTCTCTCTCTCTCTCT | 992 |
| DB | 250 | eValTyrPheLeuPheThrProTyrAsnIleValIleLeuLeuAsnThrPheGlnG | 270 |
| QY | 993 | ATTCTTTGGCTGAAATATCTACAGTCTCTACAGTCTGACAGTCTGACAGTCTG | 1052 |
| DB | 270 | ubPhePheLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnVal | 290 |
| QY | 1053 | ACAGACTCTTGGATGACGACTCTCTCATCAACCCCATCATCTATGCTCTGCTG | 1112 |
| DB | 290 | rgIuThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheValG | 310 |
| QY | 1113 | GAACTTCAGAAATCTACTCTTAGTCTTCTTCCAAAGACATTCCTCAACGCTCT | 1172 |
| DB | 310 | uLysPheArgTyrLeuSerValPhePheArgLysHisIleThrIleArgPheCys | 330 |
| QY | 1173 | ATGCTGTCTTATTTTCCAGCAAGGCTCCGAGGAGCAAGCTCTAGTTTACACCG | 1232 |
| DB | 330 | sGlnCysProValPheTyrArgGluThrValAlaValThrSerThrAsnThrPro | 350 |
| QY | 1233 | CACCTGGGAGCAGGAATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1263 |
| DB | 350 | rThrGlyGluGlnGluValSerAlaGlyLeu | 360 |

RESULT 3

I38450

chemokine (C-C) receptor 2, splice form A - human

N;Alternate names: C-CR-2; monocyte chemoattractant protein 1 receptor; monocyte c

C;Species: Homo sapiens (man)

C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 13-Aug-1999

C;Accession: I38450

R;Charo, I.F.; Myers, S.J.; Herman, A.; Franci, C.; Connolly, A.J.; Coughlin, S.R.

A;Title: Molecular cloning and functional expression of two monocyte chemoattractant

A;Reference number: A53477; MUID:94195821; PMID:8146186

A;Accession: I38450

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-374 <RES>

A;Cross-references: EMBL:U03882; NID:g472555; PIDN:AAA19119.1; PID:g472556

C;Genetics:

A;Gene: GDB:CMKBR2

A;Cross-references: GDB:337364; OMIM:601267

A;Map position: 3p21-3p21

C;Superfamily: vertebrate rhodopsin

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmemb

F;44-68/Domain: transmembrane #status predicted <TM1>

F;79-99/Domain: transmembrane #status predicted <TM2>

F;115-136/Domain: transmembrane #status predicted <TM3>

F;154-178/Domain: transmembrane #status predicted <TM4>

F;208-226/Domain: transmembrane #status predicted <TM5>

F;244-265/Domain: transmembrane #status predicted <TM6>

F;292-309/Domain: transmembrane #status predicted <TM7>

F;14/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;52-277,113-190/disulfide bonds: #status predicted

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 6.07e-97 | Length: | 374 |
| Score: | 1191.00 | Matches: | 234 |
| Percent Similarity: | 83.86% | Conservative: | 31 |
| Best Local Similarity: | 74.05% | Mismatches: | 34 |
| Query Match: | 45.32% | Indels: | 17 |
| DB: | 2 | Gaps: | 4 |

US-09-938-703-3 (1-1442) x I38450 (1-374)

| | | | |
|----|-----|---|-----|
| QY | 234 | ACAAGATGATATCAAGTCCAGTCCCAATCATCATCAATATATTATACATCGGAG | 293 |
| DB | 14 | AsnGluSerGlyGluGluValThrThr---PhePheAspTyrAspTyr-----GlyAla | 30 |
| QY | 294 | CCCTGCCAAATATCATGTGACAGCAATTCGACCCCGCTCTCTCTCTCTCTCTCT | 353 |

```

Db      31  ProCysHisLysPheAspValLysGlnIleValAlaGlnLeuLeuProLeuTyrSer 50
QY      354  CTGGTGTTCATCTTTGGTGGTGGCAATGCTGGTCAATCCATCCATGATAAAGC 413
Db      51  LeuValPheIlePheGlyPheValGlyAsnMetLeuValValLeuLeuLeuAsnCys 70
QY      414  AAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTT 473
Db      71  LysLysLeuLysCysLeuThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeu 90
QY      474  TCTCTCTTACTGTCCTCTGGGCTGCTACATGCTGCGCCGCAAGTGGAGCTTTGGAAT 533
Db      91  PheLeuIleThrLeuProLeuTrpAlaHisSerAlaAlaAsnGlnTrpValPheGlyAsn 110
QY      534  ACAATGTGTCAACTGTGACAGGCTCTATTTATAGCTCTCTCTCTGGAATCTCTTC 593
Db      111  AlaMetCysLysLeuPheThrGlyLeuTyrHisIleGlyTyrPheGlyGlyIlePhe 130
QY      594  ATCATCTCTCGAATCATGATAGTACCTGGCTGCTGCTCCATGCTGTGTTGCTTAAAA 653
Db      131  IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLys 150
QY      654  GCAGAGCGGTCACCTTTGGGCTGTCAGAGTGTGATCATCTGGTGGTGGCTGTGTT 713
Db      151  AlaArgThrValThrPheGlyValValThrSerValIleThrTrpLeuValAlaValPhe 170
QY      714  GCGTCTCCGCAAGATCATCTTACCAGATCTCAAAAGAGGCTCTTACATACACCTGC 773
Db      171  AlaSerValProGlyIleIlePheThrLysCysGlnLysGluAspSerValTyrValCys 190
QY      774  AGC-----TCTCATTTTCCATCATTAAGATAGTATCATCTTGGG 812
Db      191  GlyProTyrPheProArgGlyTrpAsnAsn-PheHisThrIleMetArgAsnIleLeuG 210
QY      813  GCTGCTCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 872
Db      210  YleuValLeuProLeuLeuIleMetValIleCysTyrSerGlyIleLeuLysThrLeu 230
QY      873  TCGGTGTCGAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCCATCATCATGAT 932
Db      230  uArgCysArgAsnGluLysLysArgHisArgAlaValArgValIlePheThrIleMet 250
QY      933  TGTATTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 992
Db      250  eValTyrPheLeuPheThrTrpProTyrAsnIleValIleLeuLeuAsnThrPheGln 270
QY      993  ATTCTTGGCTGAATATTGCTAGCTCTAACAGGTGGACCAAGCTATGACAGGTGAC 1052
Db      270  uPhePheGlyLeuSerAsnCysGluSerThrSerGlnLeuAspGlnAlaThrGlnVal 290
QY      1053  AGAGACTCTGGATGAGCACTGCTGCATCAACCCATCATCATGCTGCTGCTGCTGCTG 1112
Db      290  rGlnThrLeuGlyMetThrHisCysCysIleAsnProIleTyrAlaPheValGlyG 310
QY      1113  GAAGTTCAGAAATACCTCTTAGTCTTCTTCCAAAGACATGCTCC 1158
Db      310  uLysPheArgSerLeuPhe-----HisIleAla 319

```

RESULT 4

```

A:Title: chemokine (C-C) receptor 1 - human
N:Alternate names: C-C CKR-1; macrophage inflammatory protein-1-alpha receptor
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: A45177; 155671
R:Neote, K.; Digregorio, D.; Mak, J. Y.; Horuk, R.; Schall, T. J.
Cell 72, 415-425, 1993
A:Title: Molecular cloning, functional expression, and signaling characteristics of a C-
cell 72, 415-425, 1993
A:Reference number: A45177; MUID:93161416; PMID:7679326
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA

```

```

A:Residues: 1-355 <NEO>
A:Cross-references: GB:L10918; NID:g292416; PIDN:AAA36543.1; PID:g292417
A:Experimental source: HL60 cells
A:Note: sequence extracted from NCBI backbone (NCBIP:124876)
R:Gao, J.
J. Exp. Med. 177, 1421-1427, 1993
A:Title: Structure and functional expression of the human macrophage inflammatory 1 a
A:Reference number: 155671; MUID:93240122; PMID:7683036
A:Accession: 155671
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-355 <RES>
A:Cross-references: GB:L10918; NID:g292416; PIDN:AAA36543.1; PID:g292417
C:Genetics:
A:Gene: GDB:CMKRL1; CMKR-1
A:Cross-references: GDB:138446; OMIM:601159
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: disulfide bond; G protein-coupled receptor; glycoprotein; phosphoprotein;
F:71-91/Domain: transmembrane #status predicted <TM1>
F:108-129/Domain: transmembrane #status predicted <TM2>
F:147-171/Domain: transmembrane #status predicted <TM3>
F:205-223/Domain: transmembrane #status predicted <TM4>
F:240-264/Domain: transmembrane #status predicted <TM5>
F:288-305/Domain: transmembrane #status predicted <TM6>
F:5/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:24-273.106-183/Disulfide bonds: #status predicted
F:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted

```

Alignment Scores:

```

Pred. No.: 2, 2e-80 Length: 355
Score: 1003.50 Matches: 192
Percent Similarity: 71.31% Conservative: 59
Best Local Similarity: 54.55% Mismatches: 83
Query Match: 38.18% Indels: 18
DB: 2 Gaps: 5

```

```

US-09-938-703-3 (1-1442) x A45177 (1-355)

```

```

QY      243  GATTATCAAGTGTCAATCCATCTATGACATCAATTATTATACATGGAGCCCTCCCAA 302
Db      9  AspTyrAspThrThr-----GluPheAspTyrGlyAspAlaThrProCysGln 25
QY      303  AAAATCAATGTAAGCAAAATCGCAGCCGCTCTGCTCCGCTCTACTCTACTCTGCTG 362
Db      26  LysValAsnGluArgAlaPheGlyAlaGlnLeuLeuProLeuTyrSerLeuValPhe 45
QY      363  ATCTTTGGTTTGGGCAACATGCTGCTCATCTCTCATCTGATAAAGGCTG 422
Db      46  ValIleGlyLeuValGlyAsnIleLeuValValLeuValGlnTyrLysArgLeu 65
QY      423  AAGAGCATGCTGACATCTACTGCTCAACCTGGCCATCTGACCTGTTTCTCTT 482
Db      66  LysAsnMetThrSerIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPhe 85
QY      483  ACTGTCCCTTCTGGGCTCACTATGCTGCCGCC---CAGTGGGACTTTGGAAATACA 539
Db      86  ThrLeuProPheTrpIleAspTyrLysLeuLysAspTrpValPheGlyAspAlaMet 105
QY      540  TGTCACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTGGAATCTTCTTCATCAT 599
Db      106  CysLysIleLeuSerGlyPheTyrTyrThrGlyLeuTyrSerGluIlePhePheIle 125
QY      600  CTCTTGACAAATGATAGTACCTGCTGCTCCATGCTGCTGTTGCTTTAAAGCCAGG 659
Db      126  LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuArgAla 145
QY      660  ACGGTCACTTTGGGTTGGTGAAGTGTATGATCATTGGTGGTGGTGGTGGTGGTGG 719
Db      146  ThrValThrPheGlyValIleThrSerIleIleThrPheAlaLeuAlaIleLeuAs 165
QY      720  CTCCCGAAGTCACTCTTTTACCAGATCTCAAAAGAGGTCTTCTTACCTGACGCTCT 779

```


A:Molecule type: DNA

A:Residues: 1-355 <R>

A:Cross-references: EMBL:U28404; NID:g881547; PIDN:AAA89153.1; PID:g881548

C:Superfamily: vertebrate rhodopsin

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 6 346-76 | Length: | 355 |
| Score: | 953.00 | Matches: | 185 |
| Percent Similarity: | 69.97% | Conservative: | 55 |
| Best Local Similarity: | 53.94% | Mismatches: | 88 |
| Query Match: | 36.26% | Indels: | 15 |
| DB: | 2 | Gaps: | 4 |

US-09-938-703-3 (1-1442) x I49339 (1-355)

```
QY 270 GACATCAATTATATATACATCGAGCCCTCCCAAAATCAATGTGAAGCAATCGCACC 329
   ::: ::::: ||| ||||| ||||| ||||| |||||
Db 15 GluPheAspTyrGlyAspSerThrProCysGlnLysThrAlaValargAlaPheGlyAla 34
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 330 GCGCTCTGCTCCGCTCTACTACTGCTGTTCACTTTGGTTTGGTGGGCAACATGCTG 389
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 35 GlyLeuLeuProLeuTyrSerLeuValPheIleIleGlyValValGlyAsnValLeu 54
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 390 GTCACTCTCATCTCTGATAAAGCTGCAAAAGCTGAAGAGCATGCTGACATCTACTCTCTC 449
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 55 ValIleLeuValLeuMetGlnHisArgArgLeuGlnSerMetThrSerIleTyrLeuPhe 74
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 450 AACTGCGCATCTGCTGCTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 509
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75 AsnLeuAlaValSerAspLeuValPheLeuPheThrLeuProPheThrPheAspTyrLys 94
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 510 GCGGCC--CAGTGGGACTTGGAAATACAAATGCTCACTCTTGCAGGCGCTCTATTCT 566
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 LeuLysAspAspIlePheGlyAspAlaMetCysLysLeuLeuSerGlyPheTyr 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 567 ATAGGCTCTCTTCTGGAATCTTCTCATCTCTCTCTGACAAATGATAGTACTGCTGCT 626
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 115 LeuGlyLeuTyrSerGluIlePhePheIleIleLeuLeuThrIleAspArgTyrLeuAla 134
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 627 GTCTCTCATCTGCTGTTTGGTCTTAAAGCAGGAGGCTACCTTTGGGTGGTGGCAAGT 686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 IleValHisAlaValPheAlaLeuArgAlaArgThrValThrLeuGlyIleThrSer 154
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 687 GTGATCACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 IleIleThrTrpAlaLeuAlaIleLeuAlaSerMetProAlaLeuTyrPhePheIleAla 174
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 747 CAAGAAGAGTCTTCATACCTGAGCTCTCATTTTCCATAC----- 791
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 175 GlnTrpGluPheThrHisArgThrCysSerProHisPheProTyrLysSerLeuLysGln 194
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 792 -----ATPAAGATAGTCACTCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 833
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 TrpLysArgPheIleAlaLeuLysLeuAsnLeuLeuGlyLeuIleLeuProLeuLeuVa 214
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 834 CATGTCATCTGCTACTCGGAATCTCTAAATACTCTGCTGGTGGTGGTGGTGGTGGTGGTGG 893
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 214 ILeIleCysTyrAlaGlyIleIleArgIleLeuLeuArg---ArgProSerGluLys 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 894 GAGCAGCAGGCTGTGAGGCTTACTTCCACCATGATGTTGTTATTTTCTCTCTCTGGC 953
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 sLysValLysAlaValArgLeuIlePheAlaIleThrLeuLeuPhePheLeuLeuTrp 253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 954 TCCCTACACATGTCCTTCTCTGACACCTCTCCAGCAATCTTTGGCGCTCATATTC 1013
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 IProTyrAsnLeuSerValPheValSerAlaPheGlnAspValLeuPheThrAsnGlnCy 273
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1014 CAGTAGCTCTPACAGGTTGGACCAAGCTATGCAAGTGACAGACTCTTGGGATGACGCA 1073
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273 sGlnGlnSerLysHisLeuAspLeuAlaMetGlnValThrGluValIleAlaTyrThrHi 293
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1074 CTGCTGATCAACCCCATCTATCTATGCTTTTCTGGGGAGAGCTTCAGAACTACCTCTT 1133
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 293 sCysValAsnProIleIleTyrValPheValGlyGluArgPheTrpLysTyrLeuAr 313
QY 1134 AGTCTCTTCCAAAGACACATGCAACAACTCTGCAAACTGCTTCTATTTTCCACA 1193
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 gGlnLeuPheGlnArgHisValAlaIleProLeuAlaLysTrpLeuProPheLeuSerVa 333
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1194 AGAGCTCCCGAGCGAGCAAGCTCAGTTTACACCGATCCACTGGGAGCAGGAATATC 1253
   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 LaspGlnLeuGluArgThrSerSerIle---SerProSerThrGlyGluHisLuleuSe 352
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1254 TGTGGGC 1260
   | |||
Db 352 rAlagly 354

RESULT 7
G02436
chemokine (C-C) receptor 3 - human
N:Alternate names: C-C CKR-3
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 04-Mar-2000
C:Accession: G02436; A57237
R:Ponath, P.D.
submitted to the EMBL Data Library, February 1996
A:Reference number: H01272
A:Accession: G02436
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-355 <P>
A:Cross-references: EMBL:U49727; NID:g1477560; PIDN:AA09726.1; PID:g1477561
R:Combadiere, C.; Ahuja, S.K.; Murphy, P.M.
J. Biol. Chem. 270, 16491-16494, 1995
A:Title: Cloning and functional expression of a human eosinophil CC chemokine recepto
A:Reference number: A57237; MUID:95348056; PMID:7622448
A:Accession: A57237
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-106, N', 108-275, S', 277-280, R', 282-355 <C>
A:Cross-references: GB:U28694; NID:g1199579; PIDN:AA050469.1; PID:g1199580
A>Note: the translated sequence in GenBank entry HSU28694, release 113.0, PIDN:AA0504
C:Genetics:
A:Gene: GDB:CMK9R3
A:Cross-references: GDB:579624; OMIM:601268
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:36-60/Domain: transmembrane #status predicted <TM1>
F:71-91/Domain: transmembrane #status predicted <TM2>
F:108-129/Domain: transmembrane #status predicted <TM3>
F:147-171/Domain: transmembrane #status predicted <TM4>
F:205-223/Domain: transmembrane #status predicted <TM5>
F:240-261/Domain: transmembrane #status predicted <TM6>
F:288-305/Domain: transmembrane #status predicted <TM7>
F:24-273, 106-183/Disulfide bonds: #status predicted
F:345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicte
```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 4 88e-72 | Length: | 355 |
| Score: | 909.00 | Matches: | 176 |
| Percent Similarity: | 70.87% | Conservative: | 60 |
| Best Local Similarity: | 52.85% | Mismatches: | 82 |
| Query Match: | 34.59% | Indels: | 15 |
| DB: | 2 | Gaps: | 4 |

US-09-938-703-3 (1-1442) x G02436 (1-355)

```
QY 297 TGCCAAAAATCAATGTGAAGCAAAATCGAGCCGCTCTGCTGCTACTACTGCT 356
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24 CysGluLysAlaAspThrArgAlaLeuMetAlaGlnPheValProLeuTyrSerLeu 43
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 357 GGTTCATCTTTGTTTGGTGGCAACATGCTGCTCATCTCTCATCTGATAACTGCAAA 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44 ValPheThrValGlyLeuLeuGlyAsnValValValValIleLeuLeuLysTyrArg 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

QY 417 AGGCTGAAGCAGCATGCTACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTC 476
Db 64 ArgLeuArgLleMetThrAnIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPhe 83
QY 477 CTTCTTACGTCCCTCTCTGGCTCACTACTGCTGCCGCC---CAGTGGCACTTTGGAAT 533
Db 84 LeuValThrLeuProPheThrPheHisTyrValArgGlyHisAsnTrpValPheGlyHis 103
QY 534 ACAATGTCAACTCTGACAGGGCTATTTATAGGCTCTCTCTCGGAATCTCTTC 593
Db 104 GlyMetCysLeuLeuSerGlyPheTyrHisThrGlyLeuTyrSerGluIlePhePhe 123
QY 594 ATCATCTCTGACAACTGATAGTACTGCTGCTGCTCCACCTGCTGTTTGGCTTAA 653
Db 124 IleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuArg 143
QY 654 GCCAGGACGGTCACTTTGGGGTGTGACAGTGTGATCACTGTTGGTGGTGGTGT 713
Db 144 AlaArgThrValThrPheGlyValIleThrSerIleValThrTrpGlyLeuAlaVal 163
QY 714 GCGTCTCTCCAGGAATCACTTTACCATGCTCAAAAGAGGCTCTTCATTACCTGC 773
Db 164 AlaAlaLeuProGluPheIlePheTyrGluThrGluGluLeuPheGluGluThrLeu 183
QY 774 AGCTCT-----CATTTCCATACATTAAGAT 800
Db 184 SerAlaLeuTyrProGluAspThrValTyrSerTrpArgHis-PheHisThrLeuArg 203
QY 801 AGTCATCTTGGGGTGGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
Db 203 tThrIlePheCysLeuValLeuProLeuLeuValMetAlaIleCysTyrThrGlyIle 223
QY 861 AAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
Db 223 eLysThrLeuLeuArgCysProSer---LysLysLysTyrLysAlaIleArgLeuIle 242
QY 921 CACCATCATGATGTTTATTTCTCTCTGGGCTCCCTACACATGCTCTCTCTGAA 980
Db 242 eValIleMetAlaValPheIlePheThrTrpTyrAsnValAlaIleLeuLeu 262
QY 981 CACCTTCCAGGAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1040
Db 262 rSerTyrGlnSerIleLeuPheGlyAsnAspCysGluArgThrLysHisLeuAspLeu 282
QY 1041 TATGCAAGTGCAGAGACTCTTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCT 1100
Db 282 lMetLeuValThrGluValIleAlaTyrSerHisCysMetAsnProValIleTyrAl 302
QY 1101 CTTTGTGGGAGAGTTCAGAACTACTCTTGTAGTCTTCTTCCAAAAGCACATTC 1160
Db 302 aPheValGlyGluArgPheArgLysTyrLeuArgHisPhePheHisArgHisLeuLeu 322
QY 1161 ACGGTCTCAGAAAGCTGTTCTATTTCCACAGAGGCTCCGAGGAGGAGGAGG 1220
Db 322 tHisLeuGlyArgTyrIleProPheLeuProSerGluLysLeuGluArgThrSerVa 342
QY 1221 TTACACCCGATCCACTGGGAGCAGGAATATCTGTG 1257
Db 342 l---SerProSerThrAlaGluProGluLeuSerIle 353

```

RESULT 8

```

A57160
chemokine (C-C) receptor 4 - human
N;Alternate names: C-C CKR-4
C;Species: Homo sapiens (man)
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
R;Power, C.A.; Meyer, A.; Nemeth, K.; Bacon, K.B.; Hoogewerf, A.J.; Proudfoot, A.E.I.;
J. Biol. Chem. 270, 19495-19500, 1995
A;Title: Molecular cloning and functional expression of a novel CC chemokine receptor cd
A;Reference number: A57160; MUID:95370289; PMID:7642634
A;Accession: A57160
A;Status: preliminary; not compared with conceptual translation

```

```

A:Molecule type: mRNA
A:Residues: 1-360 <POM>
A:Cross-references: GB:X85740; NID:g1370103; PIDN:CAA59743.1; PID:g971452
A>Note: source clone K5-5
C:Genetics:
A:Gene: GDB:CKR4
A:Cross-references: GDB:677463
A:Map position: 3p21-3p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:40-65/Domain: transmembrane #status predicted <TM>
F:76-97/Domain: transmembrane #status predicted <TM>
F:112-133/Domain: transmembrane #status predicted <TM>
F:151-175/Domain: transmembrane #status predicted <TM>
F:208-226/Domain: transmembrane #status predicted <TM>
F:243-264/Domain: transmembrane #status predicted <TM>
F:291-308/Domain: transmembrane #status predicted <TM>
F:29-276/110-187/Disulfide bonds: #status predicted
F:72-350/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predi
F:145/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predi
F:183,194/Binding site: carboxylate (Asn) (covalent) #status predicted
F:321/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predi

```

Alignment Scores:

```

Pred. No.: 1,72e-70 Length: 360
Score: 891.50 Matches: 174
Percent Similarity: 67.99% Conservative: 66
Best Local Similarity: 49.29% Mismatches: 93
Query Match: 33.92% Indels: 20
DB: Gaps: 6

```

US-09-938-703-3 (1-1442) x A57160 (1-360)

```

QY 243 GATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATCGAG----- 293
Db 8 AspThrThrLeuAspGluSerIleTyrSer---AsnTyrTyrLeuTyrGluSerIlePro 26
QY 294 ---CCCTGCCCAAAAATCAATGTCGAACAATGCGAGCCGCCCTCTGCTCGCTAC 350
Db 27 LysProCysThrLysGluGlyIleLysAlaPheGlyGluLeuPheLeuProLeuTyr 46
QY 351 TCATCTGTTCTATCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 410
Db 47 SerLeuValPheValPheGlyLeuLeuGlyAsnSerValValValLeuValLeuPheLys 66
QY 411 TCGAAAGGCTGAAGAGCATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTG 470
Db 67 TyrLysArgLeuArgSerMetThrAspValTyrLeuLeuAsnLeuAlaIleSerAspLeu 86
QY 471 TTTTTCCTCTTACTGCTCCCTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
Db 87 LeuPheValPheSerLeuProPheTrpGlyTyrTyrAlaAlaAspGlnTrpValPheGly 106
QY 531 AATCAATGTGTCAACTCTTGACAGGCTCTATTTATAGGCTCTTCTCTGGAATCTTC 590
Db 107 LeuGlyLeuCysLysMetIleSerTrpMetTyrLeuValGlyPheTyrSerGlyIlePhe 126
QY 591 TTCATCATCTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATG 650
Db 127 PheValMetLeuMetSerIleAspArgTyrLeuAlaIleValHisAlaValPheSerLeu 146
QY 651 AAAGCCAGGAGCTACCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 710
Db 147 ArgAlaThrThrLeuThrTyrGlyValIleThrSerLeuAlaThrTrpSerValAlaVal 166
QY 711 TTTGCGTCTCTCCAGGAATCATTTTACCATGCTCAAAAGAGGCTCTTCATTACACC 770
Db 167 PheAlaSerLeuProGlyPheLeuPheSerThrCysTyrThrGluArgAsnHisThrTyr 186
QY 771 TCGAGC-----TCTCATTTTCCATCATTAAGAT 800
Db 187 CysLysThrLysTyrSerLeuAsnSerThrThrTrpLys-ValLeuSerSerLeuGluI 206

```

QY 375 GTGGCAGACATGCTGTGCATCTCTCATCTCGATAAATGCCAAAGCTGAAGCATGACT 434
Db 55 PheGlyAsnSerValValLeuValLeuPheLysTyrLysArgLeuLysSerMetThr 74
QY 435 GACATCTACCTGCTCAACCTGCCATCTCTGACCCTGTTTTTCCTTTCTTACTGTCCTCCCTTC 494
Db 75 AspValTyrLeuLeuAsnLeuAlaIleSerAspLeuPheValLeuSerLeuProPhe 94
QY 495 TGGGCTCACTATGTCGCCCGCAGCTGGAGCTTTGGAAATACAATGTGTCAAATCTTGACA 554
Db 95 TrpGlyTyrTyrAlaLaalaspGlnTrpValPheGlyLeuGlyLeuCysylsilevalser 114
QY 555 GGCTCTATTATATAGCTTCTCTGGAATCTTCTCATPCATCTCTGACAAATCGAT 614
Db 115 TrpMetTyrLeuValGlyPheTyrSerGlyIlePhePheIleMetLeuMetSerIleasp 134
QY 615 AGTACCTGGCTGCTGCCATCTGCTGTGTGTATAAGCACAGGAGCTGACCTTTGGG 674
Db 135 ArgTyrLeuAlaIleValHisAlavalPheSerLeuLysAlaargThrLeuThrTyrGly 154
QY 675 GTGGTGACAAGTGTGATCACTTGGGTGGTGTGTGTGCGTCTCTCCACAGGAATCATC 734
Db 155 ValIleThrSerLeuIleThrTrpSerValAlavalPheAlaSerLeuProGlyLeuLeu 174
QY 735 TTITACCATCTCAAAAAGAAGCTTTCATTACACTCCAGCTCTCATTTTC----- 786
Db 175 PheSerThrCysTyrThrGluHisAsnHisThrTyrCysLysThrGlnTyrSerValas 194
QY 787 -----CATACATTAAGATATGATCATCTGCGGCGTGTGCTGCC 824
Db 194 nSerThrThrTrpLysValLeuSerSerLeuLeuLeuasnValLeuGlyLeuLeuIleR 214
QY 825 GCTGTCTGTATGGTCATCTGCTACTCGGGAATCTAAAACTCTCTCGGTGTGCGAA 884
Db 214 OleuGlyIleMetLeuPheTyrTyrSerMetIleIleArgThrLeuGlnHisCysLysas 234
QY 885 TGAGAAGAAGAGGCAGAGGCTGTGAGGCTTATCTTACCATCATCATGTGTATTTCCT 944
Db 234 nGluLysLys---AsnArgAlaValArgMetIlePheGlyValValLeuPheLeuGl 253
QY 945 CTTCTGGGCTCCCTCAACATGTCTCTCTGCTGAACACCTTCACAGAAATCTTTGGCT 1004
Db 253 yPheThrPhr-ProTyrAsnValValLeuPheLeuGluThrLeuValGluLeuGluValle 273
QY 1005 GAATATTGCACTAGCTCTACAGGTGGACACACTATGCGGTGACAGAGACTCTTGG 1064
Db 273 uGlnaspCysThrLeuGluArgTyrLeuAspTyrAlaIleGlnAlaThrGluThrLeuGl 293
QY 1065 GATGACGCACTGCTGCATCAACCCCATCATCATGCTTTGTCGGGAGAGTTTCAGAA 1124
Db 293 yPheIleHisCysLysLeuAsnProValIleTyrPhePheLeuGlyGluLysPheArgly 313
QY 1125 CPACTCTTAGTCTTCTTCCAAAAGCACATATGCCAAACGC--TTCTGCAAAATGCTGTC 1181
Db 313 sfyrlleThrGlnLeuPheArgThrCysArgGlyProLeuValLeuCysLysHiscysas 333
QY 1182 TATTTTCCACAGAGGCTCCCGAGCGACAGCTAGTTTACACCCGATCCACTGGGA 1241
Db 333 pPheLeuGlnValTyrSerAlaAspMetSerSerSerTyrThrGlnSerThrValas 353
QY 1242 GCAGGAA 1248
Db 353 phiIsasp 355

RESULT 10
S55394
G protein-coupled receptor E1 - equine herpesvirus 2
CjSpecies: equine herpesvirus 2
CjDate: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 13-Aug-1999
CjAccession: S55594
RjReford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995

| | | | | |
|---|---|------|---|------------|
| D | b | 311 | nCysAlaLeuSerSerAsnLeuAspMetalaLeuLeuIleThrLysThrValaLaTyTh | 331 |
| | | | : : : : | |
| Q | y | 1071 | GCAGTGTGATCAACCCCATCATCTATCGCTTGTCGGGGAGAAGITCAGAACA | TACT 1130 |
| | | | : | |
| D | b | 331 | rHiscCysIleAsnProValIleTyrAlaPheValIGlyGLuLysPheArgHisIe | 351 |
| | | | : | |
| Q | y | 1131 | CTTAGTCTCTTCCAAAGACATTCGCAAGCGTTCTGCAATGCTGTCTATTTC | A 1190 |
| | | | | |
| D | b | 351 | uYrHisPheHeHisThrTyrValAlaIleTyrLeuCysLysTyrIleProPheLeu | S 371 |
| | | | | |
| Q | y | 1191 | G-----CAAGAGCGCTCCGAGCGA | 1209 |
| | | | : : | |
| D | b | 371 | rGIyAspGlyGlUGlyLysGluGlyProThrArg | 382 |
| | | | | |
| | | | RESULT 11 | |
| | | | I49340 | |
| | | | MIP-1 alpha receptor like-1 - mouse | |
| | | | C/Species: Mus musculus (house mouse) | |
| | | | C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999 | |
| | | | C/Accession: I49340 | |
| | | | R/Gao, J.L.; Murphy, P.M. | |
| | | | J. Biol. Chem. 270, 17494-17501, 1995 | |
| | | | A>Title: Cloning and differential tissue-specific expression of three mouse bet | |
| | | | A/Reference number: I49339; MUID:95340546; PMID:7542241 | |
| | | | A/Accession: I49340 | |
| | | | A>Status: preliminary; translated from GB/EMBL/DBJ | |
| | | | A/Molecule type: DNA | |
| | | | A/Residues: 1-356 <RES> | |
| | | | A/Cross-references: EMBL:U28405; NID:g881549; PIDN:AAA89154,1; PID:g881550 | |
| | | | C/Superfamily: vertebrate rhodopsin | |
| | | | Alignment Scores: | |
| | | | Pred. No.: 2,73e-60 Length: 356 | |
| | | | Score: 776.00 Matches: 158 | |
| | | | Percent Similarity: 63.56% | |
| | | | Best Local Similarity: 44.63% | |
| | | | Query Match: 29.53% | |
| | | | Indels: 19 | |
| | | | Gaps: 5 | |
| | | | DB: | |
| | | | US-09-938-703-3 (1-1442) x I49340 (1-356) | |
| Q | y | 252 | GTGTCAAGTCGAATCATATGACATC-----AATTATTATACATCGAGCCCTGC | 299 |
| | | | : | |
| D | b | 6 | ValThrGluProSerTyrAsnThrValAlaLysAsnAspPheMetSerGlyPheLeu | Cys 25 |
| | | | | |
| Q | y | 300 | CAAAAAACAATGTGAGCAATCGAGCCGCCCTCTCGCTCCGCTCTACTCTACTGTG | 359 |
| | | | | |
| D | b | 26 | PheSerIleAsnValArgAlaPheGlyIleThrValProThrProLeuTyrSerLeu | Val 45 |
| | | | | |
| Q | y | 360 | TTCATCTTTGGTTTTGTGGCAACATCGTGGTCATCTCATCTCATATAAATCAAAGG | 419 |
| | | | | |
| D | b | 46 | PheIleIleGlyValIleGlyHisValLeuValValLeuIleGlnHisLysArg | 65 |
| | | | | |
| Q | y | 420 | CTGAGACGATGACTGACATCTACTGCTCAACTGGCCACTCTGCACTGTTTTCTCT | 479 |
| | | | | |
| D | b | 66 | LeuArgAsnMetThrSerIleTyrLeuPheAsnLeuAlaIleSerAspLeuValPhe | Leu 85 |
| | | | | |
| Q | y | 480 | CTTACTGTCCCCTCTGGGCTCATA---GTCGCCGCCAGTGGGACTTTGGAATACA | 536 |
| | | | | |
| D | b | 86 | SerThrLeuProPheTrpValAspTyrIleMethysGlyAspTrpIlePheGlyAsn | Ala 105 |
| | | | | |
| Q | y | 537 | ATGTGTCAACTCTTGACAGGGCTATTATTAGAGCTTCTCTGGAATCTTCTCATC | 596 |
| | | | | |
| D | b | 106 | MetCysLysPheValSerGlyPheTyrTyrLeuGlyLeuTyrSerAspMetPhe | PheIle 125 |
| | | | | |
| Q | y | 597 | ATCCTCTGTGACANPGAVAGGTACTGGGCTGCGFCACGCTGTGTTTGCCTTTAAAG | CC 656 |
| | | | | |
| D | b | 126 | ThrLeuLeuThrIleAspArgTyrLeuAlaValHisValValPheAlaLeuArg | Ala 145 |
| | | | | |
| Q | y | 657 | AGAGCGTCCACTTTGGGTGGTGACAGTGTGATCATCTGGGTGGTGGCTGTGTTG | CG 716 |
| | | | | |
| D | b | 146 | ArgThrValThrPheGlyIleIleSerSerIleIleThrTrpValLeuAlaLeu | Val 165 |
| | | | | |

C:Genetics:

A:Gene: GDB:CMKBR8; CMKBR12; TER1; OKR-L1

A:Cross-References: GDB:6053733; OMIM:601834

A:Map position: 3p21-3p21

C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; transmembrane protein

F:36-63/Domain: transmembrane #status predicted <TM1>

F:73-94/Domain: transmembrane #status predicted <TM2>

F:108-129/Domain: transmembrane #status predicted <TM3>

F:147-171/Domain: transmembrane #status predicted <TM4>

F:200-222/Domain: transmembrane #status predicted <TM5>

F:239-260/Domain: transmembrane #status predicted <TM6>

F:281-304/Domain: transmembrane #status predicted <TM7>

Alignment Scores:

| Pred. No.: | 5,43e-55 | Length: | 355 |
|------------------------|----------|---------------|-----|
| Score: | 716.00 | Matches: | 144 |
| Percent Similarity: | 59.72% | Conservative: | 71 |
| Best Local Similarity: | 40.00% | Mismatches: | 114 |
| Query Match: | 27.25% | Indels: | 31 |
| DB: | 2 | Gaps: | 7 |

US-09-938-703-3 (1-1442) x JCS067 (1-355)

| | | | |
|----|-----|---|----------|
| QY | 240 | ATGGATTATCAAGTGCACGATCCCATCTATGACATC---AATTATTAT | -----284 |
| Db | 1 | MetAspTyrThrLeuAspLeuSerValThrThrValThrAspTyrTyrProaspIle | 20 |
| QY | 285 | ACATCGGAGCCCTGCCAAAATCAAGTGAACAAATGCGAGCCGCCCTCTCGCTCCG | 344 |
| Db | 21 | PheserSerProCysaspAlaGluLeuIle6LInThrAsnGlyLysLeuLeuAlaVal | 40 |
| QY | 345 | CTCTACTCACCTGGTTCATCTTTGGTTTGGGCAACATGCTGGTCACTCTCATCTGTG | 404 |
| Db | 41 | PhetYrCysLeuLeuPheValPheSerLeuLeuGlyAsnSerLeuValIleLeuValIle | 60 |
| QY | 405 | ATAACTCAAAAGCGTGAAGAGATGACTGACATCTACTGCTCAACTGGCCATCTCT | 464 |
| Db | 61 | ValValCysLysLysLeuArgSerIleThrAspValTyrLeuLeuAsnLeuAlaLeuSer | 80 |
| QY | 465 | GACCTGTTTTTCTCTTACTGTCCTCTGCGCTCTACTGCTGCGCCAGTGGGAC | 524 |
| Db | 81 | AspLeuLeuPheValPheSerPheProPheGlnThrTyrTyrLeuLeuAspGlnTrpVal | 100 |
| QY | 525 | TTTGGAAATACAAATGTCATCTTGTACAGAGGCTCTATTTATAGCTCTCTCTCTGGA | 584 |
| Db | 101 | PhesIyThrValMetCysLysValValSerGlyPheTyrTyrIleGlyPheYrSerSer | 120 |
| QY | 585 | ATCTTCTTCATCATCTCTCTGACATCGATAGTACTGCTGCTGCTCATGCTGCTGTTT | 644 |
| Db | 121 | MetPhePheIleThrLeuMetSerValAspArgTyrLeuAlaValValHisAlaValTyr | 140 |
| QY | 645 | GCTTTAAAGCCAGGAGGCTACCTTTGGGGTGGTGACAGTGTGATCTTGGTGGTG | 704 |
| Db | 141 | AlaLeuLysValArgThrIleArgMetGlyThrThrLeuCysLeuAlaValTripleuThr | 160 |
| QY | 705 | GCTGTGTTTGGCTCTCCGAGGATCATCTTTACCAGA---TCTCAAAAGAGGCTTT | 761 |
| Db | 161 | AlaIleMetAlaThrIleProLeuLeuValPheTyrGlnValAlaSerGluaspIyAla | 180 |
| QY | 762 | CATT-----ACACTGCAGCTCTCATTTTCTTCATCATATT | 794 |
| Db | 181 | -LeuGlnCysTyr-SerPheTyrAsnGlnGlnThrLeuLysTrpLysIlePheThrAsnPh | 200 |
| QY | 795 | AAAGATACTCATCTTTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG | 854 |
| Db | 200 | elysMetAsnIleLeuGlyLeuIleProPheThrIlePheMetPheCysTyrIleLy | 220 |
| QY | 855 | AATCTCTAAAACACTCTGCTCGGTGTCGAAATGAGAGAGAGGACACAGGCTGTCAGGCT | 914 |
| Db | 220 | sIleLeuHisGlnLeuLysArgCysGlnAsnHisAsnLysThr---LysAlaIleArgLe | 239 |
| QY | 915 | TATCTTCACCATCATGATTGTTTATTTTCTCTCTCTGGGCTCCCTACACATGTCCTTCT | 974 |

| | | | |
|------|----|--|------|
| 104 | Db | LysLeuThrThrAlaPhePhePheIleGlyPhePheGlyIlePhePheIleThrVal | 123 |
| 603 | QY | CTGCAACATGATAGGTACTGGTCTGCTGCCATCCTGTTGCTTTAAAGCAGGACG | 662 |
| 124 | Db | IIeSerIleAspArgTyrLeuAlaIleValLeuAlaAsnSerMetAsnArgThr | 143 |
| 663 | QY | GTACACTTTGGGTGGTGACAAAGTGATGATCACTTGGGTGGTGGTGTGTTGGCTCTCTC | 722 |
| 144 | Db | ValGlnHisGlyValThrIleSerLeuGlyValTrpAlaAlaIleLeuValAlaSer | 163 |
| 723 | QY | CCAGGATCATCTTTACC---AGATCTCAAAAGAAGGTCTTT---CAATTACACTGCAGC | 776 |
| 164 | Db | ProGlnPheMetPheThrLysArgLysAspAsnGluCysLeuGlyAspTyr-ProGluVa | 183 |
| 777 | QY | TTCTC-----ATTTTCCATACATTAAAGATAGTCATCTGGGCTGGTCTCT | 821 |
| 183 | Db | IleuGlnGluIleTrpProValLeuArgAsnSerGluValAsnIleuGlyPheValLe | 203 |
| 822 | QY | GCGGCTGCTTGTCATGGTCATCTGCTACTCGGGAATCTTAAAACTCGTCTGGGTGCG | 881 |
| 203 | Db | uProLeuLeuIleMetSerPheCysTyrPheArgIleValArgThrLeuPheSerCysLy | 223 |
| 882 | QY | AAATGACAGACAGACGACGAGGCTGTGAGGTTATCTTCACATCATGATGTTTATTT | 941 |
| 223 | Db | sAsnArgLysLysAla--ArgAlaIleArgLeuIleLeuLeuValValValPhePh | 242 |
| 942 | QY | TTCTCTTCTGGGCTCCCTACACACATCTGCCTCTCTCTGACACCTTCCAGAAATCTTTGG | 1001 |
| 242 | Db | eLeuPheThrThrProTyrAsnIleValIlePheLeuGluThrLeuLysPheTyrAsnPh | 262 |
| 1002 | QY | CTGAAATATTTGCAGTAGTCTTAACAGGTTGGACCAAGCTATGCAGGTGCACAGACTCT | 1061 |
| 262 | Db | ePheProSerCysGlyMetLysArgAspLeuArgTrpAlaLeuSerValThrGluThrVa | 282 |
| 1062 | QY | TGGGATCAGCCACTGCTCCATCAACCCCATCATGTATGCTTTGTCGGGAGAGTTGAG | 1121 |
| 282 | Db | IAlaPheSerHisCysCysLeuAsnProPheIleTyrAlaPheAlaGlyLysPheAr | 302 |
| 1122 | QY | AAACTACTCTTAGTCTTCTTCCAAAGACGATTCGCAACGCTCTGCAATGCTGTC | 1181 |
| 302 | Db | gArgTyrLeu-----ArgHisLeu-----TyrAsnLysCysLeuAl | 314 |
| 1182 | QY | TATTTTC-----CAGCAAGAGGCTCC | 1202 |
| 314 | Db | aValLeuCysGlyArgProValHisAlaGlyPheSerThrGluSerGlnArgSerArgL | 334 |
| 1203 | QY | CGACGGCAGACGCTCAGTTTACACCGATCCACTGGGGACGAGGAATATCTGTG | 1257 |
| 334 | Db | nAspSerIleLeuSerLeuThrHisTyrThrSerGluGlyGlySerLeu | 352 |

RESULT 14
JC4304
orphan G protein-coupled receptor - human
N:Alternate names: V28 protein
C:Species: Homo sapiens (man)
C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 19-May-2000
C:Accession: JC4304
R:RAPoport, C.J.; Schweickart, V.L.; Eddy Jr., R.L.; Shows, T.B.; Gray, P.W.
Gene 163, 295-299, 1995
A:Title: The orphan G-protein-coupled receptor-encoding gene V28 is closely related to the human G-protein-coupled receptor-encoding gene V28
A:Reference number: JC4304, MUID:96011651; PMID:7590284
A:Accession: JC4304
A:Molecule type: mRNA
A:Residues: 1-355 <RAP>
A:Cross-references: GB:U0350; NID:g665580; PIDN:AAA91783.1; PID:g665581
A:Experimental source: peripheral blood mononuclear cell
C:Comment: This protein is a cell-surface receptor which recognizes extracellular ligands and is involved in the regulation of many immune and homeostatic responses.
C:Genetics:
A:Gene: V28
A:Map position: 3pter-p21
C:Superfamily: vertebrate rhodopsin
C:Keywords: G-protein-coupled receptor; lymphokine; transmembrane protein


```
Db 199 -PheLeuThrLeuLysMetAsnIleSerValLeuValLeuProLeuPheIlePheThrPh 218
QY 843 CTGCTACTCGGAACTCTTAAACTCTGCTCGGTGTCGAAATGAGAGAGAGACACAG 902
Db 218 eLeuTyValGlnMetArgLysThrLeu-----ArgPheArgGluGlnArgTyrSe 235
QY 903 GGCTGTGAGGCTTATCTTCACCATCATGATTGTTATTCTCTCTCTGGGCTCCCTACAA 962
Db 235 rLeuPheLysLeuValPheAlaIleMetValPheLeuLeuMetIrpAlaProTyrAs 255
QY 963 CATTGCTCTCTCTGAACACCTTCCAGGAATCTTTGGCCTGAAATAATTGCAGTAGCTC 1022
Db 255 nIleAlaPhePheLeuSerThrPheLysGluHisPheSerLeuSerAspCysLysSerSe 275
QY 1023 TAACAGGTTGGACCAAGCTATCCAGGTGACAGAGACTCTTGGGATGACCGACTGCTGCAT 1082
Db 275 rTyrAsnLeuAspLysSerValHisIleThrLysLeuIleAlaThrThrHisCysCysI 295
QY 1083 CAACCCATCATCTATGCTTTGTCGGGAGAGATTTCAGAACTACTCTTACTTCTTCTT 1142
Db 295 eAsnProLeuLeuTyAlaPheLeuAspGlyThrPheSerLysTyrIleu----- 311
QY 1143 CCAGAGACATTGCCAAACGCTTCTGCAAAATGCTGTCTATTATTTCCAGCAAGAGCTCC 1202
Db 312 -----CysArgCysPheHisLeu---ArgSerAsnThrPr 322
QY 1203 CGAGCGAGCAAGCTCAGTTTACACCGATCCACTGGGAGCAGCAA 1248
Db 322 oLeuGlnProArgGlyGlnSerAlaGlnGlyThrSerArgGluGlu 337
```

Search completed: June 3, 2003, 18:51:41
Job time : 65.6953 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:39:54 ; Search time 23.4701 Seconds
(without alignments)
5096.611 Million cell updates/sec

Title: US-09-938-703-3
Perfect score: 2628
Sequence: 1 GAATCCCCACACAGACCA.....AGTAGATTAGATCCGAATC 1442

Scoring table:

| | |
|---------------------------|--|
| BLOSUM62 | |
| Xgapop 10.0 , Xgapext 0.5 | |
| Ygapop 10.0 , Ygapext 0.5 | |
| Fgapop 6.0 , Fgapext 7.0 | |
| Delop 6.0 , Delext 7.0 | |

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-O/cyn2/1/USPTO.spool/US0938703/runat_03062003.161406.22808/app_query.fasta_1.1870
-DB-SwissProt_40 -OFMT=FASTAN -SUFFIX=rs -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938703.ecgn.1.1.38-brunat_03062003.161406.22808 -NCPU=6 -ICPU=3
-NO_WMAP
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------------|---------------------|
| 1 | 1745.5 | 66.4 | 352 | 1 CKR5_HUMAN | P51681 homo sapien |
| 2 | 1739.5 | 66.2 | 352 | 1 CKR5_PANTR | P56440 pan troglod |
| 3 | 1734.5 | 66.0 | 352 | 1 CKR5_PONPY | O97881 pongo pygma |
| 4 | 1729.5 | 65.8 | 352 | 1 CKR5_GORGO | P56439 gorilla gor |
| 5 | 1725.5 | 65.7 | 352 | 1 CKR5_TRAPH | O97879 trachypithe |
| 6 | 1720.5 | 65.5 | 352 | 1 CKR5_PAPHA | P56441 papio hamad |
| 7 | 1720.5 | 65.5 | 352 | 1 CKR5_TRAFR | O97878 trachypithe |
| 8 | 1719.5 | 65.4 | 352 | 1 CKR5_PYGBI | O97880 pygathrix b |
| 9 | 1719.5 | 65.4 | 352 | 1 CKR5_PYGNE | O97882 pygathrix n |
| 10 | 1718.5 | 65.4 | 352 | 1 CKR5_HYLLE | O97883 hylobates l |
| 11 | 1717.5 | 65.4 | 352 | 1 CKR5_MACMU | P79436 macaca mula |
| 12 | 1707.5 | 65.0 | 352 | 1 CKR5_CERTO | O62743 cercopithec |
| 13 | 1694.5 | 64.5 | 352 | 1 CKR5_CERAE | P56493 cercopithec |
| 14 | 1475.5 | 56.1 | 354 | 1 CKR5_RAT | O08556 rattus norv |
| 15 | 1473.5 | 56.1 | 354 | 1 CKR5_MOUSE | P51682 mus muscullu |
| 16 | 1328.5 | 50.6 | 373 | 1 CKR2_MOUSE | P51683 mus muscullu |
| 17 | 1326.5 | 50.5 | 373 | 1 CKR2_RAT | O55193 rattus norv |
| 18 | 1315.5 | 50.1 | 360 | 1 CKR2_MACMU | O18793 macaca mula |

| | | | | | |
|----|--------|------|-----|--------------|---------------------|
| 19 | 1191 | 45.3 | 374 | 1 CKR2_HUMAN | P41597 homo sapien |
| 20 | 1003.5 | 38.2 | 355 | 1 CKR1_HUMAN | P32246 homo sapien |
| 21 | 976 | 37.1 | 359 | 1 CKR3_MOUSE | P51678 mus muscullu |
| 22 | 974 | 37.1 | 359 | 1 CKR3_RAT | O54814 rattus norv |
| 23 | 950 | 36.1 | 355 | 1 CKR1_MOUSE | P51675 mus muscullu |
| 24 | 949.5 | 36.1 | 355 | 1 CKR3_MACMU | P56482 macaca mula |
| 25 | 912 | 34.7 | 355 | 1 CKR3_HUMAN | P51677 homo sapien |
| 26 | 902 | 34.3 | 355 | 1 CKR3_MACMU | P56483 macaca mula |
| 27 | 898.5 | 34.2 | 358 | 1 CKR3_CAVPO | O92423 cavia porce |
| 28 | 891.5 | 33.9 | 360 | 1 CKR4_HUMAN | P51679 homo sapien |
| 29 | 884 | 33.6 | 355 | 1 CKR3_CERAE | P56492 cercopithec |
| 30 | 878.5 | 33.4 | 360 | 1 CKR4_MOUSE | P51680 mus muscullu |
| 31 | 776 | 29.5 | 356 | 1 CKR7_MOUSE | P51676 mus muscullu |
| 32 | 716 | 27.2 | 355 | 1 CKR8_HUMAN | P51685 homo sapien |
| 33 | 709.5 | 27.0 | 356 | 1 CKR8_MACMU | O97665 macaca mula |
| 34 | 708.5 | 27.0 | 353 | 1 CKR8_MOUSE | P56484 mus muscullu |
| 35 | 688.5 | 26.2 | 354 | 1 C3X1_RAT | P35411 rattus norv |
| 36 | 669.5 | 25.5 | 355 | 1 C3X1_MOUSE | O92089 mus muscullu |
| 37 | 644 | 24.5 | 355 | 1 C3X1_HUMAN | P49238 homo sapien |
| 38 | 582.5 | 22.2 | 357 | 1 CKR9_HUMAN | P51686 homo sapien |
| 39 | 573.5 | 21.8 | 378 | 1 CKR7_HUMAN | P32248 homo sapien |
| 40 | 566.5 | 21.6 | 369 | 1 CKR9_MOUSE | O92077 mus muscullu |
| 41 | 556.5 | 21.2 | 378 | 1 CKR2_MOUSE | P47774 mus muscullu |
| 42 | 556.5 | 21.2 | 384 | 1 CKD6_HUMAN | O00590 homo sapien |
| 43 | 555 | 21.1 | 378 | 1 CKD6_MOUSE | O08707 mus muscullu |
| 44 | 554 | 21.1 | 382 | 1 CKD6_RAT | O09027 rattus norv |
| 45 | 552.5 | 21.0 | 374 | 1 CKR6_HUMAN | P51684 homo sapien |

ALIGNMENTS

RESULT 1

CKR5_HUMAN

| | | | | |
|----|--|---|------|--------|
| ID | CKR5_HUMAN | STANDARD: | PRT: | 352 AA |
| AC | P51681 | 014692; 014693; 014695; 014696; 014697; 014698; 014699; | | |
| AC | 014700; 014701; 014702; 014703; 014704; 014705; 014706; 014707; | | | |
| AC | 014708; 015358; 09UPA4; | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | |
| DT | 15-OCT-2002 (Rel. 41, Last annotation update) | | | |
| DE | C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5) | | | |
| DE | (HIV-1 fusion co-receptor) (CHEMR13) (CD195 antigen). | | | |
| GN | CCR5 OR CKR5 | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE-96241590; PubMed-8639485; | | | |
| RA | Samson M., Labbe O., Mollereau C., Vassart G., Parmentier M.; | | | |
| RT | "Molecular cloning and functional expression of a new human | | | |
| RT | CC-chemokine receptor gene." | | | |
| RL | Biochemistry 35:3362-3367(1996). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE-96291862; PubMed-8663314; | | | |
| RA | Raport C.J., Gosling J., Schweichart V.L., Gray P.W., Charo I.F.; | | | |
| RT | "Molecular cloning and functional characterization of a novel human | | | |
| RT | CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha." | | | |
| RL | J. Biol. Chem. 271:17161-17166(1996). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE-96295970; PubMed-8699119; | | | |
| RA | Combadiere C., Ahuja S.K., Tiffany H.L., Murphy P.M.; | | | |
| RT | "Cloning and functional expression of CC CKR5, a human monocyte CC | | | |
| RT | chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and | | | |
| RT | RANTES." | | | |
| RL | J. Leukoc. Biol. 60:147-152(1996). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | McCombie W.R., Willson R., Chen E., Gibbs R., Zuo L., Johnson D., | | | |
| RA | Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K., | | | |

FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).
 FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION.
 FT MOD_RES 10 10 SULFATION.
 FT MOD_RES 14 14 SULFATION.
 FT MOD_RES 15 15 SULFATION.
 FT VARIANT 10 10 /FTID=VAR_003481.
 FT VARIANT 29 29 Y -> D (IN INCCR5-71A).
 FT VARIANT 31 31 A -> S (IN DBSNP:1800939).
 FT VARIANT 31 31 /FTID=VAR_011839.
 FT VARIANT 31 31 R -> H (IN INCCR5-72A).

Alignment Scores:
 Pred. No.: 5,25e-131 Length: 352
 Score: 1745.50 Matches: 340
 Percent Similarity: 96.60% Conservative: 1
 Best Local Similarity: 96.32% Mismatches: 0
 Query Match: 66.42% Indels: 12
 DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CRR5_HUMAN (1-352)

QY 240 ATGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGC 299
 Db 1 MetAspArgGlnValSerSerProIleThrAspIleAsnThrThrSerGluProCys 20
 QY 300 CAAAATCAATGTGAAGCAAAATCGCAGCCGCCCTCTGCTCCGCTCTACTCACTGGG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuThrSerLeuVal 40
 QY 360 TTCACTTTGGTTTGGGCACACATGCTGTCATCTCTCTGCTGCTGCTGCTGCTGCTG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACTTACCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 Db 61 LeuLysSerMetThrAspIleThrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTACTGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 Db 81 LeuThrValProPheThrAlaHisTyraAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCACTCTTCACAGGCTCTATTATATAGCTCTCTCTGCTGCTGCTGCTGCTGCTGCT 599
 Db 101 CysGlnLeuLeuThrGlyLeuThrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGACATCGATGATGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaValValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTACCTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCACGAATCATCTTTACAGATCTCAAAAGAAGTCTTCATTACACCTGCGAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisThrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAAGATAGTTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValI 200
 QY 807 CTGGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTCTCGGTGTGCAATGAG 926
 Db 220 fLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240

QY 927 CAGATGTTGTTTATTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTGAATAATTCAGTACGTCTAACAGGTTGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGACATCTTTGGGATGACGACATCTGTCATCAACCCCATCATCTATGCTTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAAAATGCTGTCTATTCTTCCAGAGAGGCTCCCGAGGAGGAGGAGGAGGAGGAGG 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CGATCCACTGGGAGCAGGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 2
 CRR5_PANTR
 ID CRR5_PANTR STANDARD; PRT; 352 AA.
 AC P56440; O02778;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
 GN CCR5 OR CXCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97268687; PubMed=9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Pelper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RT Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zimmerman P.A., Buckler-White A., Alkhatib G.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98022612; PubMed=9359654;
 RA Zhang L., Carruthers C.D., He T., Huang Y., Cao Y., Wang G., Hahn B.,
 RA Ho D.D.;
 RT "HIV type 1 subtypes, coreceptor usage, and CCR5 polymorphism.";
 RL AIDS Res. Hum. Retroviruses 13:1357-1366(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97426118; PubMed=9282822;
 RA Zacharova V., Zachar V., Goustin A.S.;
 RT "Sequence of chemokine receptor gene CCR5 in chimpanzees, a natural
 RT HIV type 1 host.";
 RL AIDS Res. Hum. Retroviruses 13:1159-1161(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98090115; PubMed=9430250;
 RA Pretet J.-L., Zerbib A., Girard M., Guillet J.-G., Butor C.;
 RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1.";
 RL AIDS Res. Hum. Retroviruses 13:1583-1587(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE

RX MEDLINE-9726687; PubMed-9108095;
 RA Edinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF005659; AAB62553.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm1.1; 1
 DR PRINTS; PRO0237; GPCRHOPOPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 FT DOMAIN 1 30
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58
 FT DOMAIN 59 68
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89
 FT DOMAIN 90 102
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124
 FT DOMAIN 125 141
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166
 FT DOMAIN 167 198
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218
 FT DOMAIN 219 235
 FT CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260
 FT DOMAIN 261 277
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301
 FT DOMAIN 302 352
 FT CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178
 FT MOD_RES 3 3
 FT MOD_RES 10 10
 FT MOD_RES 14 14
 FT MOD_RES 15 15
 FT SULFATION (BY SIMILARITY).
 FT SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40515 MW; D0E6FCB9FE5EAC84 CRC64;
 Alignment Scores:
 Pred. No.: 9,78e-130 Length: 352
 Score: 1729.50 Matches: 336
 Percent Similarity: 96.03% Conservative: 3
 Best Local Similarity: 95.18% Mismatches: 2
 Query Match: 65.81% Indels: 12
 DB: 1 Gaps: 1
 US-09-938-703-3 (1-1442) x CKR5_GORGO (1-352)
 QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTG 299
 Db 1 MetaspYrGlnValSerSerProThrTyAspIleAspTyThrSerGluProCys 20
 QY 300 CAAATAATCATGTAGCAATATCGAGCCGCTCTCGCTCTACTACTACTGCTG 359
 Db 21 GlnLysThrAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTySerLeuVal 40
 QY 360 TTCATCTCTGTTTGGGCAACATGCTGCTCATCTCTCATCTGATACATGCAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuIleAsnCysIysArg 60

QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCT 479
 Db 61 LeuLysSerMetThrAspIleTyLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCTCTGCGGCTCCTATGCTGCGCCAGTGGGACTTTGGAAATACAATG 539
 Db 81 LeuThrValProPheThrAlaHisTyAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyPheIleCysPheSerGlyIlePhePheIle 120
 QY 600 CTCTGACAACTGATAGTACCTGGCTGCTGCTCCATCTGCTGTTTGTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgTyLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACTTTGGGTGGTGACAGTGTGATCATCTGGGTGGGTGGTGGTGGTGGTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCGCCGAATCATCTTTACAGATCTCAAAAAGAGGTCTTCTATTACCTGACCTGACCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyThrCysSerSer 180
 QY 780 CATTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTySerGlnTyGlnPheTrpLysAsnPheGlnThrLeu-LysIleVal 200
 QY 807 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTySerGlyIleLeuLys 220
 QY 867 TCTGCTCGGTGTCGAAATGAGAAGAGAGAGACAGGCTGTGAGGCTTATCTTCCACAT 926
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATTTCTTTATTTCTCTCTGCGCTCCCTACACATGCTCTCTCTGACACCTT 986
 Db 240 eMetIleValTyPheLeuPheThrPalaProTyAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTGAATAATTGAGTAGCTCTAACAGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMet 280
 QY 1047 GGTGACAGACTCTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyAlaPheVa 300
 QY 1107 CGGGAGAGTTCAGAACTACTCTTAGTCTTCTTCCAAAGACACATTCGCAAGCGCTT 1166
 Db 300 lGlyGluLysPheArgAsnTyLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATCTGTTCTTATTTTCCAGAGAGGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyThr 340
 QY 1227 CCGATCCACTGGGAGGAGGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGlnGlnIleSerValGlyLeu 352
 RESULT 5
 ID CKR5_TRAPH STANDARD; PRT; 352 AA.
 AC 097879;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CCR-5) (CC-CR-5) (CCR-5) (CCR5).
 GN CCR5 OR CCR5R5.
 OS Trachypithecus phayrei (Phayre's leaf monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecoidea; Papio.
 CC NCBI_TaxID=9557, 9555;
 RN [1]

RN SEQUENCE FROM N.A.
 RC SPECIES=P.hamadryas;
 RX MEDLINE=97268687; PubMed=9108095;
 RA Eddinger A.L., Amedee A., Miller K., Doranz B.J., Endres M.,
 RA Sharron M., Samson M., Lu Z.H., Clements J.E., Murphy-Corb M.,
 RA Peiper S.C., Parmentier M., Broder C.C., Doms R.W.;
 RT "Differential utilization of CCR5 by macrophage and T cell tropic
 RT simian immunodeficiency virus strains.";
 RN Proc. Natl. Acad. Sci. U.S.A. 94:4005-4010(1997).
 RN [2]

RN SEQUENCE FROM N.A.
 RC SPECIES=P.hamadryas;
 RX MEDLINE=99210133; PubMed=10195758;
 RA Sakkena N.K., Wang B., Novembre F.J., Bolton W., Smit T.K., Lal R.B.;
 RT "Species-specific changes in the CCR5 gene from African and Asian
 RT nonhuman primates.";
 RN AIDS Res. Hum. Retroviruses 15:479-483(1999).
 RN [3]

RN SEQUENCE FROM N.A.
 RC SPECIES=P.hamadryas;
 RX Benton P.A., Timanus D.K., Shearer M.H., Lee D.R., Kennedy R.C.;
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AF005658; AB62552.1;
 CC EMBL: AF105287; AAD20556.1;
 CC EMBL: AF105288; AAD20557.1;
 CC EMBL: AF105289; AAD20558.1;
 CC EMBL: AF105290; AAD20559.1;
 CC EMBL: AF023452; AAC63830.1;
 CC InterPro: IPR000276; GPCR_Rhodan.
 CC Pfam: PF00001; 7tm.1; 1.
 CC PRINTS: PR00237; GPCRHOOPS.
 CC PROSITE: PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
 CC PROSITE: PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 CC DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 31 58 1 (POTENTIAL).
 CC DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 69 89 2 (POTENTIAL).
 CC DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 103 124 3 (POTENTIAL).
 CC DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 142 166 4 (POTENTIAL).
 CC DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 199 218 5 (POTENTIAL).
 CC DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 236 260 6 (POTENTIAL).
 CC DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 278 301 7 (POTENTIAL).
 CC DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 CC DISULFID 101 178 SULFATION (BY SIMILARITY).
 CC MOD_RES 3 3 SULFATION (BY SIMILARITY).
 CC MOD_RES 10 10 SULFATION (BY SIMILARITY).

FT MOD_RES 14 14 SULEFATION (BY SIMILARITY).
 FT MOD_RES 15 15 SULEFATION (BY SIMILARITY).
 FT CARBOHYD 268 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 352 AA; 40489 MW; 5E15049BA1FE8B2 CRC64;
 Alignment Scores:
 Pred. No.: 5,07e-129 Length: 352
 Score: 1720.50 Matches: 333
 Percent Similarity: 96.03% Conservative: 6
 Best Local Similarity: 94.33% Mismatches: 2
 Query Match: 65.47% Indels: 12
 DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CKR5_PAPHA (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCATGACATCAATTTATATACATGCGGCCCTGC 299
 Db 1 MetaspTyrGlnValSerSerProThrTyrAspIleAspTyrThrSerGluProCys 20
 QY 300 CAAAATAATCAATGTGAAGCAAAATCGACGCGCTCTCTGCTCGCTCTACTACTACTGCTG 359
 Db 21 GlnLysIleasnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTTGGTTTGTGGGCAACATGCTGGTCATCTCATCTGATGAAATCAATG 419
 Db 41 PheIlePheGlyPheValGlyAsnIleValValLeuLeuLeuLeuLeuLeuLeuLeuVal 60
 QY 420 CTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCTGGGCTCACTGCTGCGCCGACAGTGGGACTTTGAAATCAATG 539
 Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTGTGACGGCTCTATTTATAGCTTCTTCTCTGGAATCTTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIle 120
 QY 600 CTCTGCAATCATGATGATCTTACCAAGATCTCAAAAAGAGCTCTTCACTTACCTGAGCTCT 779
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTGTCACCTTTGGGGTGTGACAGGTGATCATCTGGGTGGTGGCTGTGTTGGCTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCCAGGAATCATCTTTTACCAAGATCTCAAAAAGAGCTCTTCACTTACCTGAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValII 200
 QY 807 CTTGGGCTGCTGCTGCGCTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTGGGTGCGAAATGAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuLeuIlePheThrII 240
 QY 927 CATGATTTGTTATTTCTCTCTGGGCTCCCTACACATTTGCTCTTCTCTGAGACCTT 986
 Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTGGAATATGAGTAGTCTTAACAGGTGGAGAGAGAGAGAGAGAGAG 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGI 280
 QY 1047 GGTGACAGAGACTCTTGGGATGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 1047 GGTGACAGAGACTCTTGGGATGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT


```

|||||
280 nValThrGluThrLeuGlyMetThrHisCysSylleasnProIleIleTyAlaPheVa 300
|||||
QY 1107 CGGGAGAGAGTTGAGAACTACCTCTTCTTCTTCCAAAGCAGCATTCGCAACGCTT 1166
|||||
Db 300 IGIGLULysPheArgAsnTyLeuLeuValPheGlnLysHisIleAlaLysArgPh 320
|||||
QY 1167 CTVCAAAATGCTGTTCTATTTCACGACGAGGCTCCGAGGAGGAGCTAGTTTACAC 1226
|||||
Db 320 ecYALysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerValTyTh 340
|||||
QY 1227 CGGATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTG 1263
|||||
Db 340 rArgSerThrGlyGluGlnGluThrSerValGlyLeu 352

RESULT 8
CKR5_PYGBI
ID CKR5_PYGBI STANDARD; PRT; 352 AA.
AC 097880;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-C chemokine receptor type 5 (C-CR-5) (CCR-5) (CCR5).
GN CKR5 OR CKR5.
OS Pygathrix bieti (Black snub-nosed monkey) (Rhinopithecus bieti).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Pygathrix.
OC NCBI_TaxID=61621;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99416438; PubMed=10486970;
RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
RT "Sequence evolution of the CKR5 chemokine receptor gene in primates.";
RL Mol. Biol. Evol. 16:1145-1154(1999).
CC -I- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are NO restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF075445; AAD19857.1; -.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; glycoprotein; Sulfation.
CC DOMAIN 1 30
CC TRANSMEM 31 58
CC DOMAIN 59 68
CC TRANSMEM 69 89
CC TRANSMEM 90 102
CC TRANSMEM 103 124
CC DOMAIN 125 141
CC TRANSMEM 142 166
CC TRANSMEM 167 198
CC TRANSMEM 199 218
CC TRANSMEM 219 235
CC TRANSMEM 236 260
CC TRANSMEM 261 277
CC TRANSMEM 278 301
CC TRANSMEM 302 352

```

```

FT DISULFID 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
SQ SEQUENCE 352 AA; 40585 MW; 4366FI42730F938F CRC64;

Alignment Scores:
Pred. No.: 6,08e-129 Length: 352
Score: 1719.50 Matches: 332
Percent Similarity: 96.03% Conservativeness: 7
Best Local Similarity: 94.05% Mismatches: 2
Query Match: 65.43% Indels: 12
DB: 1 Gaps: 1

US-09-938-703-3 (1-1442) x CKR5_PYGBI (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATACATGGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrTyrSerGluProCys 20
QY 300 CAAAATATCAATGTGAAGCAATCGCAGCCGCTCTCTGCTCGCTCTACTCACTGGTG 359
Db 21 GlnLysValasnValysGlnIleAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGTTTGTGGGCAACATGCTGGTCATCTCTCTCTGATGATGATGATGATG 419
Db 41 PheIlePheGlyPheValGlyAsnIleValValValLeuLeuLeuLeuLeuLeuLeuLeu 60
QY 420 CTGAAGAGCATGACTGACATCTACTCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTCTCTGGCTCACTATGCTGGCCGCGCCAGTGGAGTTCGAAATACATG 539
Db 81 LeuThrValProPheTrpAlaHisTyrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTGTACAGGCTCTATTTATAGGCTTCTTCTCTGGAAATCTTCTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCTGCAATCATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValAlaValPheAlaLeuLysAlaArg 140
QY 660 AGGTCACCTTTGGGTGGTGTCACCAAGTGTGATCATCTGGGTGGGTGGGTGGGTGGGTCT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCAGGAATCATCTTACCATCTCAAAAGAGGCTCTTCTTATCTTACACCTGACGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerSer 180
QY 780 CATTTCCTCATAC-----ATTAAGATAGTGCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeuLysIleValIle 200
QY 807 CTTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTTCTGGTGTCCAAATGAGAGAGAGGACACAGGCTGCTGAGGCTTATCTTACCAT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuLeuPheThrIle 240
QY 927 CATGATTGTTTATTTCTCTCTGGGCTCCCTACACATTTGCTCTCTCTCTCTCTCTCTCT 986
Db 240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuLeuLeuTh 260
QY 987 CCAGGAATCTTTCGCTGGAATATTGAGTAGTCTTAACAGGTTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSerSer 280

```


QY 1047 GGTGACAGAGACTCTGGGATGACGCACTGCTGCATCAACCCATCATCTATGCTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleThrAlaPheVa 300
 QY 1107 CGGGAGAAGTTCAGAACTACTCTTAGTCTCTTCCAAAGACACATTCGCAAGCGTT 1166
 Db 300 lGlyGluysPheArgAsnThrLeuValPheGlnHisLeuValHisIleAlaLysArgph 320
 QY 1167 CTCGAATCTGTTCTATTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValThr 340
 QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGCGTTG 1263
 Db 340 rArgSerThrGlyGluHisGluIleSerValGlyLeu 352

RESULT 10
 CKR5_HYLLC STANDARD; PRT; 352 AA.
 AC 097883;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (CCR5).
 GN CKR5 OR CMKR5.
 OS Hylobates leucogynus (White-cheeked gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=61853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99416438; PubMed=10486970;
 RA Zhang Y.-W., Ryder O.A., Zhang Y.-P.;
 RT "Sequence evolution of the CCR5 chemokine receptor gene in primates.";
 RL Mol. Biol. Evol. 16:1145-1154(1999).
 CC -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
 CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
 CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
 CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF075451; AAD19663.1;
 CC InterPro: IPR000276; GPCR_Rhodpsn.
 CC Pfam: PF00001; 7tm_1.1;
 CC PRINTS: PR00237; GPCRHHODPSN.
 CC PROSITE: PS00237; G-PROTEIN RECP_FL1; 1.
 CC PROSITE: PS0262; G-PROTEIN RECP_FL2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation.
 KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 89 2 (POTENTIAL).
 FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 124 3 (POTENTIAL).
 FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 142 166 4 (POTENTIAL).
 FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 218 5 (POTENTIAL).
 FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 236 260 6 (POTENTIAL).
 FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 301 7 (POTENTIAL).

FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 101 178 BY SIMILARITY.
 FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
 FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
 FT MOD_RES 14 14 SULFATION (BY SIMILARITY).
 SQ SEQUENCE 352 AA; 40445 MW; 4F8B4F344CEB7C91 CRC64;
 Alignment Scores:
 Pred. No.: 7.3e-129 Length: 352
 Score: 1718.50 Matches: 334
 Percent Similarity: 95.75% Conservative: 4
 Best Local Similarity: 94.62% Mismatches: 3
 Query Match: 65.39% Indels: 12
 Db: 1 Gaps: 1
 US-09-938-703-3 (1-1442) x CKR5_HYLLC (1-352)
 QY 240 ATGGATTATCAAGTGTCAAGTCCAAATCATATGACATCAATTATTATACATCGAGCCCTGC 299
 Db 1 MetAspThrGlnValSerSerProThrThrAspIleAspThrSerGluProCys 20
 QY 300 CAAAAAATCAATGTGAGCAAAATCGAGCCCGCTCTCTGCTCGCTCTACTCTACTGGTG 359
 Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuThrSerLeuVal 40
 QY 360 TTCATCTTGGTTTGTGGCAACATCTGCTGTCATCTCTCATCTGATAAATCAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuValLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTCTCTT 479
 Db 61 LeuLysSerMetThrAspIleThrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCGCGCTCACTATGCTGCGCCAGTGGGACTTTGGAAATACATG 539
 Db 81 LeuThrValProPheThrAlaHisThrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTGACAGGCTCTATTATAGGCTCTCTCTGGAATCTCTTCTATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuThrPheIleGlyPheSerGlyIlePhePheIle 120
 QY 600 CTCCTGACAATCATGAGTACCTGCTGCTCATGCTGTGTTGTTTAAAGCCAGG 659
 Db 121 LeuLeuThrIleAspArgThrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTCACTTTGGGTGTGACAAAGTGTATCACTTGGTGGTGGTGGTGGTGGTGGT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCACGAAATCATCTTTACCATCTCAAGATCTCAAAAAGAGGTCTTCAATACACCTGACCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAAGATAGTCAT 806
 Db 181 HisPheProThrSerGlnThrGlnPheThrLysAsnPheGlnThrLeuLysIleVal 200
 QY 807 CTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysThrSerGlyIleLeuLys 220
 QY 867 TCTGCTTGGTGTGCAAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCCAT 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThr 240
 QY 927 CATGATCTTTTATTTCTCTCTGCGCTCCCTACAAATGTCTCTCTCTCTCTCTCTCT 986
 Db 240 eMetIleValThrPheLeuPheThrAlaProThrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CAGGAATCTTGGCTGAAATTAATTCAGTAGCTCTACAGGTTCGACCCAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerSerSerSerSerSerSerSer 280


```
QY 240 ATGGATTATCAAGTCTCAAGTCAATCTATGACATCAATTAATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGAAATCGCAGCCGCTCTCTGCTCCGCTCTACTCTACCTGGTG 359
Db 21 GlnIysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTGGTTTGGGCAACATCGTGGTGCATCTCATCTCATTAATCTGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnIleLeuValValLeuIleLeuIleAsnCysIleArg 60
QY 420 CTGAAGAGATGACATGACATCTACCTGCTCAACTGCGCATCTCTGACCTGTTTCCCT 479
Db 61 LeuIysSerMetThrAspIleTyrLeuLeuAsnIleAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAATACAAATG 539
Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTCGAATCTTCTTCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
Db 121 LeuLeuThrIleAspTyrLeuAlaIleValHisAlaValPheAlaLeuIleAlaArg 140
QY 660 ACGTCACTTGGTGGGTTGGGCAACATGATGATGATGATGATGATGATGATGATGATG 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCTGAGATCATCTTACCATGCTCAAAAGAGCTTCTCATTCACATGCTGAGCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerPro 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTCAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeu-LysIleValII 200
QY 807 CTTGGGCTGGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyLeuLysTh 220
QY 867 TCTGCTTCGCTGCAATGATGAGAGAGAGGACAGGCTGCTGAGCTTATCTTCACCAT 926
Db 220 rLeuLeuArgCysArgGlnGluLysArgHisArgAlaValArgLeuIlePheThrII 240
QY 927 CATGATTGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheThrTyrPheTyrAsnIleValLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTGGCTGCAATATGCAAGTAGCTCTAACAGGTGGACCAAGCTATGCA 1046
Db 260 eGlnGluPheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetGI 280
QY 1047 GGTGACAGACATCTTGGATGACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGAGAGTCTAGAACTACCTCTTGTAGTCTTCTTCCAAAGACATTCGCAACGCTT 1166
Db 300 lGlyGluIysPheArgAsnTyrLeuLeuValPhePheGlnIysHisIleAlaIysArgPh 320
QY 1167 CTGCAATGCTGTTCTATTTTCCAGAGAGGCTCCCGAGGAGGAGCTAGTTCAGTTTAC 1226
Db 320 eCysIysCysSerIlePheGlnGlnGluAlaSerGluArgAlaSerSerValTyrTh 340
QY 1227 CCGATCCACTGGGAGAGGAAATATCTGTGGGCTTG 1263
Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
RESULT 13
```

```
CKR5_CERAE
ID CKR5_CERAE STANDARD; PRT; 352 AA.
AC PS6493;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (CCR5).
GN CKR5 OR CMKR5.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=98001387; PubMed=9343222;
RA Kuhlmann S.E., Platt E.J., Kozak S.L., Kabat D.;
RT "Polymorphisms in the CKR5 genes of African green monkeys and mice
RT implicate specific amino acids in infections by simian and human
RT immunodeficiency viruses.";
RL J. Virol. 71:8642-8656(1997).
RN [2]
RN SEQUENCE FROM N.A.
RA Murayama Y., Matsunaga S., Inoue-Murayama M.;
RT "cDNA sequence of African green monkey CCR-5 chemokine receptor
RT gene.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA,
CC MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY
CC INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. MAY PLAY A ROLE
CC IN THE CONTROL OF GRANULOCYTIC LINEAGE PROLIFERATION OR
CC DIFFERENTIATION
CC !- SUBCELLULAR LOCATION: Integral membrane protein.
CC !- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U83324; AAC51795.1; -.
CC EMBL; U83325; AAC51796.1; -.
CC EMBL; AB015944; BAA31328.1; -.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm1.1;
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Sulfation;
CC Polymorphism.
KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 124 3 (POTENTIAL).
FT DOMAIN 125 141 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 142 166 4 (POTENTIAL).
FT DOMAIN 167 198 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 199 218 5 (POTENTIAL).
FT DOMAIN 219 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 260 6 (POTENTIAL).
FT DOMAIN 261 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 301 7 (POTENTIAL).
FT DOMAIN 302 352 CYTOPLASMIC (POTENTIAL).
FT DISULFD 101 178 BY SIMILARITY.
FT MOD_RES 3 3 SULFATION (BY SIMILARITY).
FT MOD_RES 10 10 SULFATION (BY SIMILARITY).
FT MOD_RES 15 15 SULFATION (BY SIMILARITY).
```


| | | | |
|-----------|-------------------|--|--------------|
| QY | 1107 | CGGGGAGAGTTCAGAAACTACCTCTTACTCTCTTCCTCCAAAGACACATGCGAAAGCGTT | 1166 |
| Db | 300 | lGyLgluylsPheArAsuTyLeuLeuValPhePheGlnLyShISlLeAlaLysArgph | 320 |
| QY | 1167 | CTGCAAAATGCTGTCTATTTCACGACGAGGCTCCCGACGAGCAAGCTCAGCTTTTACAC | 1226 |
| Db | 320 | eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerValIyTrh | 340 |
| QY | 1227 | CCGATCCACTGGGGGACGAGAAATATCTGTGGGC | 1260 |
| Db | 340 | rArgSerThrGlyGluGlnGluThrSerValGly | 351 |
| RESULT 14 | | | |
| ID | CKRS_RAT | STANDARD; | PRT; 354 AA. |
| AC | AC | 008556; | |
| DT | DT | 01-NOV-1997 (Rel. 35, Created) | |
| DT | DT | 01-NOV-1997 (Rel. 35, Last sequence update) | |
| DT | DT | 30-MAY-2000 (Rel. 39, Last annotation update) | |
| DE | DE | C-C chemokine receptor type 5 (C-C CKR-5) (CCR-5) (MIP-1 alpha receptor). | |
| DE | DE | CKRS OR CKRB5. | |
| GN | GN | Rattus norvegicus (Rat). | |
| OS | OS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OX | OX | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | |
| OX | NCBI_TaxID=10116; | | |
| RN | [1] | SEQUENCE FROM N.A. | |
| RP | RP | STRAIN-Wistar; TISSUE-Brain; | |
| RX | RX | MEDLINE-98334064; PubMed-9670989; | |
| RA | RA | Spieles O., Gournala N., Boddeke H.W.G.M., Sauter A., Fiebich B.L., | |
| RA | RA | Berger M., Gebicke-Haerter P.J.; | |
| RT | RT | "Cloning of rat hiv-1-chemokine coreceptor CKRS from microglia and | |
| RT | RT | upregulation of its mRNA in ischemic and endotoxemic rat brain." | |
| RL | RL | J. Neurosci. Res. 53:16-28(1998). | |
| RL | [2] | | |
| RP | RP | SEQUENCE FROM N.A. | |
| RP | RP | STRAIN-Sprague-Dawley; | |
| RX | RX | MEDLINE-98318173; PubMed-9655467; | |
| RA | RA | Jiang Y., Salafranca M.N., Adhikari S., Xia Y., Feng L., Sonntag M.K., | |
| RA | RA | Defiebre C.M., Pennell N.A., Streit W.J., Harrison J.K.; | |
| RT | RT | "Chemokine receptor expression in cultured glia and rat experimental | |
| RT | RT | allergic encephalomyelitis." | |
| RL | RL | J. Neuroimmunol. 86:1-12(1998). | |
| CC | CC | -1- FUNCTION: RECEPTOR FOR A C-C TYPE CHEMOKINE. BINDS TO MIP-1-ALPHA, | |
| CC | CC | MIP-1-BETA AND RANTES AND SUBSEQUENTLY TRANSDUCES A SIGNAL BY | |
| CC | CC | INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL. | |
| CC | CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. | |
| CC | CC | -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. | |
| CC | CC | | |
| CC | CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | |
| CC | CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | |
| CC | CC | the European Bioinformatics Institute. There are no restrictions on its | |
| CC | CC | use by non-profit institutions as long as its content is in no way | |
| CC | CC | modified and this statement is not removed. Usage by and for commercial | |
| CC | CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | |
| CC | CC | or send an email to license@isb-sib.ch). | |
| DR | DR | EMBL; Y12009; CAAT72737.1; - | |
| DR | DR | EMBL; U77350; AAC03243.1; - | |
| DR | DR | InterPro; IPR00276; GPCR_Rhodpsn. | |
| DR | DR | Pfam; PF00001; 7tm1.1. | |
| DR | DR | PRINTS; PR00237; GPCRHHODPSN. | |
| DR | DR | PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1. | |
| DR | DR | PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1. | |
| KW | KW | G-protein coupled receptor; Transmembrane; Glycoprotein. | |
| FT | FT | DOMAIN 1 32 | |
| FT | FT | EXTRACELLULAR (POTENTIAL). | |
| FT | FT | TRANSMEM 33 60 | |
| FT | FT | 1 (POTENTIAL). | |
| FT | FT | DOMAIN 61 70 | |
| FT | FT | CYTOPLASMIC (POTENTIAL). | |
| FT | FT | TRANSMEM 71 91 | |
| FT | FT | 2 (POTENTIAL). | |
| FT | FT | DOMAIN 92 104 | |
| FT | FT | EXTRACELLULAR (POTENTIAL). | |
| FT | FT | TRANSMEM 105 126 | |
| FT | FT | 3 (POTENTIAL). | |

FT DOMAIN 127 143 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 144 168 4 (POTENTIAL).
 FT DOMAIN 169 200 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 201 220 5 (POTENTIAL).
 FT DOMAIN 221 237 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 238 262 6 (POTENTIAL).
 FT DOMAIN 263 279 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 280 303 7 (POTENTIAL).
 FT DOMAIN 304 354 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 103 180 BY SIMILARITY.
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 354 AA; 41030 MW; 77EDB368AA4C86D CRC64;

Alignment Scores:
 Pred. No.: 1,42e-109 Length: 354
 Score: 1475.50 Matches: 283
 Percent Similarity: 86.48% Conservative: 34
 Best Local Similarity: 79.72% Mismatches: 24
 Query Match: 56.15% Indels: 14
 DB: 1 Gaps: 2

US-09-938-703-3 (1-1442) x CKR5_RAT (1-354)
 QY 240 ATGATATCAAGTGTCAAGTCCA-----ATCTATGACATCAATTATTATATACATCGAG 293
 DB 1 MetaspheglnGlySerLeuProThrTyrIleTyrAspIleAspTyrSerMetSerAla 20
 QY 294 CCCTGCCAAAATAATGTAAGCAAAATCGAGCCGCCCTCTCGCTCCGCTCTACATCA 353
 DB 21 ProCysGlnValAsnValValysGlnIleAlaAlaGlnLeuLeuProLeuTyrSer 40
 QY 354 CTGGTGTCTCATCTTTGGTGGCAACATGCTGGTCACTCTCTCATCTCTGATAACTGC 413
 DB 41 LeuValPheIlePheGlyPheValGlyAsnMetValPheLeuIleLeuSerCys 60
 QY 414 AAAAGGCTGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTT 473
 DB 61 LysIleValLeuLysSerMetThrAspIleTyrLeuPheAsnLeuAlaIleSerAspLeuLeu 80
 QY 474 TTCCTTCTACTGCTCCCTTCTGGGCTCACTATGCTGGCCAGTGGAGCTTGGAAAT 533
 DB 81 PheLeuLeuThrLeuProPheTrpAlaHisTyrAlaAlaAsnGluTrpValPheGlyAsn 100
 QY 534 ACAATGTCTCAACTCTGACAGGCTCTATTATATAGCTTCTCTCTGGAATCTCTTC 593
 DB 101 IleMetCysIleValPheThrGlyIleTyrHisIleGlyTyrPheGlyGlyIlePhePhe 120
 QY 594 ATCACTCTCTGACATCATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTTAAA 653
 DB 121 IleIleLeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaIleLys 140
 QY 654 GCAGGAGCTGACCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 713
 DB 141 AlaArgThrValAsnPheGlyValIleThrSerValValThrTrpValValAlaValPhe 160
 QY 714 GGTCTCTCCAGGAATCATCTTTACCATCTCTCAAAAGAGGCTTCTTACATACACTGC 773
 DB 161 ValSerLeuProGluIleIlePheMetArgSerGlnLysGluGlySerHisTyrThrCys 180
 QY 774 AGC-----TCTCATTTTCCATACATTAAGAT 800
 DB 181 SerProHisPheLeuHisIleGlnTyrArgPheTrpLysHis-PheGluThrLeuLysMe 200
 QY 801 AGTCATCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 860
 DB 200 TValIleLeuSerLeuIleLeuProLeuValMetValIleCysTyrSerGlyIleLe 220
 QY 861 AAAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 DB 220 uAsnThrLeuPheArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePhe 240
 QY 921 CACCATCATGATTTTATTTCTCTTCTGGGCTCCCTPACACATTTGCTCTCTCTGAA 980

Db 240 eAlaIleMetIleValTyrPheLeuPheTrpThrProTyrAsnIleValLeuLeuTh 260
 QY 981 CACCTTCCAGGAATCTTTGGCTCGAATAATTCAGTAGCTCTTAACAGGTGGACCAAGC 1040
 Db 260 rThrPheGlnGluTyrPheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAl 280
 QY 1041 TATGAGGTGACAGAGACTCTTGGATGACGACTGCTGCTCATCAACCCCATCATCTATGC 1100
 Db 280 aMetGlnValThrGlnThrLeuGlyMetThrHisCysLeuAsnProValIleTyrAl 300
 QY 1101 CTTTGTGGGGAGAAAGTTCAGAACTACCTTAGTCTCTTCCAAAGACACATTCGCAA 1160
 Db 300 aPheValGlyGluLysPheArgAsnTyrLeuSerValPhePheArgIleValVal 320
 QY 1161 ACCTTCTGCAAAATCGTGTCTATTTCCAGCAAGAGGCTCCGAGCGCAACAGTCACT 1220
 Db 320 sArgPheCysLysHisCysSerIlePheGlnValAsnProAspArgValSerSerVa 340
 QY 1221 TTACACCCGATCCACTGGGAGCAGGAATAATCTCTGTGGGCTTG 1263
 Db 340 lTyrThrArgSerThrGlyGlnGluValSerThrGlyLeu 354

RESULT 15
 CKR5_MOUSE
 ID CKR5_MOUSE STANDARD; PRT: 354 AA.
 AC P51682; Q61867; P97405; O35313; P97308; O35891;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C-C chemokine receptor type 5 (C-C CKR-5) (CC-CKR-5) (CCR-5) (MIP-1
 alpha receptor).
 GN CKR5 OR CKR5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ; TISSUE=Spleen;
 RX MEDLINE=96205938; PubMed=8631787;
 RA Boring L., Gosling J., Montecarlo F.S., Lusis A.J., Tsou C.-L.,
 Charo I.F.;
 RT "Molecular cloning and functional expression of murine JR (monocyte
 chemoattractant protein 1) and murine macrophage inflammatory protein
 1alpha receptors: evidence for two closely linked C-C chemokine
 receptors on chromosome 9.";
 RL J. Biol. Chem. 271:7551-7558(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Thymus;
 RX MEDLINE=96278910; PubMed=8662890;
 RA Meyer A., Coyle A.J., Proudfoot A.E.I., Wells T.N.C., Power C.A.;
 RT "Cloning and characterization of a novel murine macrophage
 inflammatory protein-1 alpha receptor.";
 RL J. Biol. Chem. 271:14445-14451(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Ola;
 RX Kuziel W.A., Beck M.A., Dawson T.C., Maeda N.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 and NIH Swiss; TISSUE=Liver, Kidney, and Spleen;
 RX MEDLINE=98001387; PubMed=9343222;
 RA Kumann S.E., Platt E.J., Kozak S.L., Kabat D.;
 RT "Polymorphisms in the CCR5 genes of African green monkeys and mice
 implicate specific amino acids in infections by simian and human
 immunodeficiency viruses.";
 RL J. Virol. 71:8642-8656(1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RX MEDLINE=97404635; PubMed=9261347;

| Alignment Scores: | | | |
|---|-----------|---|-------|
| Pred. No.: | 2.05e-109 | Length: | 354 |
| Score: | 1473.50 | Matches: | 281 |
| Percent Similarity: | 87.04% | Conservative: | 32 |
| Best Local Similarity: | 79.15% | Mismatches: | 28 |
| Query Match: | 56.07% | Indels: | 14 |
| DB: | 1 | Gaps: | 2 |
| US-09-938-703-3 (1-1442) x CKRS_MOUSE (1-354) | | | |
| QY | 240 | ATGGATTATCAAGTGTCCAGTCCA-----ATCTATGACATCAATATTATTATACATCGGAG | 2933 |
| Db | 1 | MetaspheGlnGlySerValProThrTyrlleIyrAspIleasnTyrglyMetserala | 20 |
| QY | 294 | CCCTGCCAAAAAATCAATGTGAAGCAAAATCGACGCCGCCCTCTCGTCCGCTCGCTACTCA | 3533 |
| Db | 21 | ProCysGlnLysIleAsnValLysGlnIleAlaGlnLeuLeuProLeuLeuTyrs | 40 |
| QY | 354 | CTGGTCTCANCATTGGTTTGTGGGCAACATGCTGGTCATCTCTCATCTCTGATAAATGCG | 4133 |
| Db | 41 | LeuValPheIlePheGlyPheValGlyAsnMetValPheLeuIleLeuIleSerCys | 60 |
| QY | 414 | AAAAGCGCTGAAGACATGATGACATCTACTCTGCTCAACTGGCCTCTCTGACCTGTTT | 4733 |
| Db | 61 | LysLysLeuLysSerValThrAspIleTyrlleuLeuAsnLeuAlaIleSerAspLeuLeu | 80 |
| QY | 474 | TTCTCTTCTACTGTCCCTCTCTGGGCTCAGTATGCTGCGGCCAGTGGGACTTTGGAAT | 5333 |
| Db | 81 | PheLeuLeuThrLeuProPheTrpAlaHisTyrlleAlaAlaAsnGluTrpIlePheGlyAsn | 1000 |
| QY | 534 | ACAATGTGCAACTTGTGACAGGCTCATTTTATAGGCTCTTCTCTGGAATCTCTTC | 5933 |
| Db | 101 | IleMetCysLysValPheThrGlyValTyrlleHisIleGlyTyrllePheGlyIlePhe | 1200 |
| QY | 594 | ATCATCTCTCTGACAATCGATAGTAGTACCTGGCTCGTCCATGCTGTGTGCTTTAAAA | 6533 |
| Db | 121 | IleIleLeuLeuThrIleAspArgTyrlleuAlaIleValHisAlaValPheAlaLeuLys | 1400 |
| QY | 654 | GCAGAGCGGTACCTTTGGGGTGTGACAGTGTGATCACTTGGGTGGGTGGTGTGTTT | 7133 |
| Db | 141 | ValArgThrValAsnPheGlyValIleThrSerValValThrTrpValValAlaValPhe | 1600 |
| QY | 714 | GCCTCTCTCCAGGAATCATCTTACCAGATCTCAAAAAGAGGTCTTCAATACACCTGC | 7733 |
| Db | 161 | AlaSerLeuProGluIleIlePheThrArgSerGlnLysGluGlyPheHisTyrlleThrCys | 1800 |
| QY | 774 | AGTCTCATTTTCCATAC-----ATTAAGAT | 8000 |
| Db | 181 | SerProHisPheProHisThrGlnTyrlleHisPheTrpLysSerPheGlnThrLeu-LysMe | 2000 |
| QY | 801 | AGTCATCTGGGGTGTCTCGCCCTGCTTGTCATGTCATCTGCTACTCGGGATCCT | 8600 |
| Db | 200 | tValIleLeuSerLeuLeuProLeuLeuValMetIleIleCysTyrlleCysTyrlleLe | 2200 |
| QY | 861 | AAAACTCTGCTTCGGTGTGCAAAATGAGAAGACAGCGCTGTGAGGCTTATCTT | 9200 |
| Db | 220 | uHisThrLeuPheArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePhe | 2400 |
| QY | 921 | CACATCATGATTTGTTATTCTCTCTGGGTCCCTACACATTTGCTCTTCTCTGAA | 9800 |
| Db | 240 | eAlaIleMetIleValTyrlleValTyrllePheLeuPheTrpThrProTyrlleValLeuLeuLeuTh | 2600 |
| QY | 981 | CACCTTCCAGGAATCTTTGGCTCAATATTCAGTACTCTTAACAGTTGGACCAAGC | 10400 |
| Db | 260 | rThrPheGlnGluPhePheGlyLeuAsnCysSerSerSerAsnArgLeuAspGlnAl | 2800 |
| QY | 1041 | TATCAGGTGTACAGAGACTCTGGGATGAGCGACTGCTCATCAACCCCATCATCTATGC | 11000 |
| Db | 280 | aMetGlnAlaThrGluThrLeuGlyMetThrHisCysCysLeuAsnProValIleTyrlle | 3000 |
| QY | 1101 | CTTTGTGGGGAGAAGTTCAAAACTACCTCTTAGTCTTCTTCCAAAGACATTTGCCAA | 11600 |

Db 300 aPheValGlyGluLysPheArgSerTyrLeuSerValPhePheArgLysHisIleVally 320
QY 1161 ACGTTCTGCAAAAGCTGTTCTATTTCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGT 1220
|||||
Db 320 sArgPheCysLysArgCysSerIlePheGlnGlnAspAsnProAspArgValSerSerVa 340
|||||
QY 1221 TTACACCGATCCACTGGGAGCAGCAATATCTGTGGGCTTC 1263
|||||
Db 340 lTyrThrArgSerThrGlyGluHisGlnValSerThrGlyLeu 354
|||||

Search completed: June 3, 2003, 18:46:01
Job time : 39.4701 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_p2p model

Run on: June 3, 2003, 18:41:24 ; Search time 103.268 Seconds
 (without alignments)
 5754.334 Million cell updates/sec

Title: US-09-938-703-3
 Perfect score: 2628
 Sequence: 1 GAATCCCGCCACAGAGCCA.....AGTAGATTAGATCCGAATTC 1442

Scoring table: BLOSUM62

| | |
|-------------|-------------|
| Xgapop 10.0 | Xgapext 0.5 |
| Ygapop 10.0 | Ygapext 0.5 |
| Fgapop 6.0 | Fgapext 7.0 |
| Delop 6.0 | Delext 7.0 |

Searched: 671580 seqs, 206047115 residues
 Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Command line parameters:

```
-MODE=framer-np.model -DBV=xlp
-Q/-cgn2_1/USPTO_spool/US0938703/runat_03062003_161407_22827/app_query.fasta_1.1870
-DB-SPTREMBL_21 -OPMT-fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0938703.ecgn_1.1_192.0runat_03062003_161407_22827 -NCP=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

| | |
|-----|------------------|
| 1: | sp_archaea: |
| 2: | sp_bacteria: |
| 3: | sp_fungi: |
| 4: | sp_human: |
| 5: | sp_invertebrate: |
| 6: | sp_mammal: |
| 7: | sp_mhc: |
| 8: | sp_organelle: |
| 9: | sp_phase: |
| 10: | sp_plant: |
| 11: | sp_rodent: |
| 12: | sp_virus: |
| 13: | sp_vertebrate: |
| 14: | sp_unclassified: |
| 15: | sp_rvirus: |
| 16: | sp_bacteriap: |
| 17: | sp_archaeap: |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|--------|--------------------|
| 1 | 1729.5 | 65.8 | 352 | 6 | Q95NC5 | Q95nc5 hylobates s |

| ID | Q95NC5 | PRELIMINARY | PRT | 352 AA |
|----|---|-------------|-----|--------|
| AC | Q95NC5 | | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Created) | | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last sequence update) | | | |
| DT | 01-MAR-2002 (TREMBLrel. 20, Last annotation update) | | | |
| DE | C-C chemokine receptor 5. | | | |
| GN | CCR5 | | | |
| OS | Hylobates syndactylus (Siamang) (Symphalangus syndactylus). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates. | | | |
| OX | NCBI_TaxID=9590; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Zhang Y., Ryder O.A., Zhang Y.; | | | |
| RT | "Sequence comparison of the CCR5 gene in primates and primate phylogeny." | | | |
| RL | Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AF177884; AAK43367.1; | | | |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. | | | |
| DR | Pfam: PF00001; 7tm1.1 | | | |
| DR | PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. | | | |
| DR | PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1. | | | |
| KW | Receptor. | | | |

ALIGNMENTS

RESULT 1

Q95NC5 ID Q95NC5 PRELIMINARY; PRT; 352 AA.

AC Q95NC5

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE C-C chemokine receptor 5.

GN CCR5

OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.

OX NCBI_TaxID=9590;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhang Y., Ryder O.A., Zhang Y.;

RT "Sequence comparison of the CCR5 gene in primates and primate phylogeny."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF177884; AAK43367.1;

DR InterPro: IPR000276; GPCR_Rhodpsn.

DR Pfam: PF00001; 7tm1.1

DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

```
SQ SEQUENCE 352 AA; 40508 MW; F4F64B3AD5AF658A CRC64;

Alignment Scores:
Pred. No.: 2,19e-162 Length: 352
Score: 1729.50 Matches: 336
Percent Similarity: 96.03% Conservative: 3
Best Local Similarity: 95.18% Mismatches: 2
Query Match: 65.81% Indels: 12
DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q95NC5 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCATCTATGACATCAATATTATATACATCGAGCCCTGC 299
Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrThrSerGluProCys 20
QY 300 CAAAAATCAATGTGAAGCAAAATCGAGCCGCTCCTCGCTCCGCTCTACTCATCTGCTG 359
Db 21 GlnLysIleAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
QY 360 TTCATCTTTGGTTTGTGGGCAACATGCTGTCATCTCATCTCATCTGATAAATGCAAAAGG 419
Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuAsnCysLysArg 60
QY 420 CTGAAGAGCATGATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479
Db 61 LeuLysSerMetThrAspIleLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
QY 480 CTTACTGTCCCTTCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
Db 81 LeuThrValProPheThrPheAlaHisTyrAlaAlaGlnTyrAspPheGlyAsnThrMet 100
QY 540 TGTCAACCTTCGACAGGCTCTATTATATAGGCTTCTCTGGAATCTTCTATCATCATC 599
Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATCAATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
QY 660 AGCGTCACCTTTGGTGTGACAAAGTGTGATCAGTGGGTGGTGGTGGTGGTGGTGGTGGT 719
Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCCTGAGAAATCAATCTTACAGATCTCAAAAGAGGTCTCTATTCACCTGCGACCTCT 779
Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGlnLysGlnLysGlnLysGlnLys 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTACAT 806
Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeuLysIleValIle 200
QY 807 CTTGGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
QY 867 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
Db 220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrIle 240
QY 927 CATGATTTGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
QY 987 CCAGGAATCTTTGGCTCAATAATGCTAGCTCTACAGGTGTGGACCAACCACTATGCA 1046
Db 260 eGlnGlnPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG1 280
QY 1047 GGTGACAGACTCTTGGGATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
Db 280 nValThrGlnThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
QY 1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTCGCAAAAGGCTT 1166
```


QY 1047 GGTGACAGACTCTGGGATGACCGACTGCTGCATCAACCCCATCATCTATGCGCTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGAGTTCAGAACTACTCTAGTCTTCTTCCAAAAGACACATTGCCAAAGCGTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTCTATTTTCCAGCAGAGGCTCCGAGGAGCAGAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 4
 Q9XS99 PRELIMINARY; PRT; 352 AA.
 ID Q9XS99; AC Q9XS99;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CC chemokine receptor 5.
 GN CCR5.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GORILLACCR;
 RX MEDLINE=99210133; PubMed=10195758;
 RA Saksena N.K., Wang B., Novembre F.J., Bolton W., Smit T., Lal R.B.;
 RT "Species-specific changes in the CCR5 gene from African and Asian
 nonhuman primates.";
 RL AIDS Res. Hum. Retroviruses 15:479-483(1999).
 DR EMBL; AF105291; AAD20560.1; -
 DR InterPro; IPR00276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40529 MW; 1B68C68FE2E7AD0 CRC64;

Alignment Scores:
 Pred. No.: 4,34e-162 Length: 352
 Score: 1726.50 Matches: 335
 Percent Similarity: 96.03% Conservative: 4
 Best Local Similarity: 94.90% Mismatches: 2
 Query Match: 65.70% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q9XS99 (1-352)

QY 240 ATGATTATCAAGTGTCAAGTCCCAATCATGACATCAATATATACATCGGAGCGCTGC 299
 Db 1 MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrThrSerGluProCys 20
 QY 300 CAAANAATCAATGTGAAGCAATGCGAGCCGCTCCGCTCCGCTCTACTACTGCTGTG 359
 Db 21 GlnLysThrAsnValLysGlnIleAlaAlaArgLeuLeuProProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTGTGTTTGTGGCAACATGCTGGTCACTCCATCCCTGATGAATGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnMetLeuValIleLeuIleLeuLeuAsnCysLysArg 60
 QY 420 CTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGACCTGTTTTCCTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTTCTGGGCTCACTATGCTGCGGCCAGGTGGGACTTTTGAATAACAATG 539

Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTACAGGCTCTATTTATAGCTTCTCTCGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCTGCAATCGATPAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACSTGTCACCTTTGGGCTGGTGGTCAAGTGCATCATCTGGTGGTGGTGGTGGTGGTCT 719
 Db 141 ThrValThrPheGlyLeuValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCAGGAATCATCTTTACCATCTCAAAAAGAGTCTTCAATACACCTGCTGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CATTTCCTATAC-----ATTAAAGATAGTTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
 QY 807 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCGCTGCTGAATGAGAAGAGGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
 Db 220 rLeuLeuArgCysArgAsnGluLysLysArgHisArgAlaValArgLeuIlePheThrI 240
 QY 927 CATCATGTTTATTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db 240 eMetIleValTyrPheLeuPheThrAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTTGGCTGCTGAATATTCAGTAGTCTTAACAGGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGACTCTTGGGATGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGAGTTCAGAACTACTCTAGTCTTCTTCCAAAAGACACATTGCCAAAGCGTT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATGCTGTCTATTTTCCAGCAGAGGCTCCGAGGAGCAGAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGCAAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 5
 Q9TW50 PRELIMINARY; PRT; 352 AA.
 ID Q9TW50; AC Q9TW50;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CC chemokine receptor type 5.
 GN CCR5.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1410;
 RX MEDLINE=99335215; PubMed=10408730;
 RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,

RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
 RT "Mutations in CCR5-coding sequences are not associated with HIV
 carrier status in African nonhuman primates.";
 RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1410;
 RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
 RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF035214; AAD44007.1;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40481 MW; 2578A0F2C07AAA65 CRC64;
 Alignment Scores:
 Pred. No.: 5,45e-162 Length: 352
 Score: 1725.50 Matches: 334
 Percent Similarity: 96.32% Conservative: 6
 Best Local Similarity: 94.62% Mismatches: 1
 Query Match: 65.66% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q9RV50 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrThrSerGluProCys 20
 QY 300 CAAAATCAATGTGAAGCAATGCGAGCCGCCCTCGCTCGCTCTACTCATCTGGTG 359
 Db 21 GlnTyrLeuValSerGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCACTCTTGTGTTTGGGCAACATGCTGCTCATCTCTCTGATCACTGCAAAAGG 419
 Db 41 PheIlePheGlyPheValGlyAsnIleLeuValIleValIleLeuIleAsnCysLysArg 60
 QY 420 CTGAAGAGCATAGTACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTTT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTACTGTCCCTTCTGGGCTCACTATGCTGCGGCCGACGTTGGAAATCAATG 539
 Db 81 LeuThrValProPheThrAlaHisTyrAlaAlaIleGlnTrpAspPheGlyAsnThrMet 100
 QY 540 TGCTCACTCTGTGAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPheSerGlyIlePhePheIleIle 120
 QY 600 CTCCTGACATCATAGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 121 LeuLeuSerIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 ACGGTCACCTTGGGCTGTGACAGTGTGATCATCTGCTGGTGGTGGTGGTGGTGGTCT 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
 QY 720 CTCCTGAGATCATCTTTACCATGATCTCAAAAGAGGCTTCTTATACCTGCAGCTCT 779
 Db 161 LeuProGlyIleIlePheThrArgSerGlnLysGluGlyLeuHisTyrThrCysSerSer 180
 QY 780 CAFTTTCATAC-----ATTAAGATAGTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrLysAsnPheGlnThrLeu-LysIleValIle 200
 QY 807 CTGGGGCTGGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
 Db 200 eLeuGlyLeuValLeuProLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220

QY 867 TCTGCTTCGGTCTCGAATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuLeuPheThrIle 240
 QY 927 CATGATTGTTATTTCTCTCTTGGGCTCCCTCAACATGTCTCTCTCTGACACACTT 986
 Db 240 eMetIleValIyrPheLeuPheThrProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATTTCTTGGCCTGAATTAATTCAGTAGCTCTAAGAGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetIle 280
 QY 1047 GTGTCAGAGACTCTTGGGATGACGACTGTGCATCAACCCCATCATCATCTCTTGT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGAGTTCAAGAACTACCTCTTACTCTCTTCCAAAAGACATGCCCCAAGCTT 1166
 Db 300 LGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAAAATGCTGTTCTATTATTTCCAGCAAGGCTCCCGAGCGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTG 1263
 Db 340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352

RESULT 6

Q95NC0 PRELIMINARY; PRT; 352 AA.
 AC Q95NC0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE C-C chemokine receptor 5.
 GN CCR5.
 OS Hylobates moloch (silvery gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Rutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 NCBI_TaxID=81572;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Ryder O.A., Zhang Y.;
 RT "Sequence comparison of the CCR5 gene in primates and primate
 RT Phylogeny.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF177899; AAK43382.1;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40436 MW; 9623CA98340CF274 CRC64;

Alignment Scores:

Pred. No.: 1,08e-161 Length: 352
 Score: 1722.50 Matches: 335
 Percent Similarity: 95.75% Conservative: 3
 Best Local Similarity: 94.90% Mismatches: 3
 Query Match: 65.54% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q95NC0 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspIleAspTyrThrSerGlyProCys 20
 QY 300 CAAAATCAATGTGAAGCAAAATCGCAGCCGCCCTCGCTCGCTCTACTCATCTGGTG 359
 Db 21 GlnLysLeuValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal 40

| | |
|----|--|
| OS | Nasalis larvatus (Proboscis monkey). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae; |
| OC | Nasalis. |
| OC | NCBI_TaxID=43780; |
| EN | [1] |
| RP | SEQUENCE FROM N.A. |
| RA | Zhang Y., Ryder O.A., Zhang Y.; |
| RT | *sequence comparison of the CCR5 gene in primates and primate |
| RT | phylogeny.*; |
| RL | Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. |
| DR | EMBL; AF177882; AA43365.1; - |
| DR | InterPro: IPR000276; GPCR_Rhodpsn. |
| DR | Pfam: PF00001; 7tm1; 1. |
| DR | PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1. |
| DR | PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1. |
| KW | Receptor. |
| SQ | SEQUENCE 352 AA; 40537 MW; 51F6F1486E35938E CRC64; |

| | | | |
|------------------------|-----------|---------------|-----|
| Alignment Scores: | | | |
| Pred. No.: | 1.36e-161 | Length: | 352 |
| Score: | 1721.50 | Matches: | 332 |
| Percent Similarity: | 96.32% | Conservative: | 8 |
| Best Local Similarity: | 94.05% | Mismatches: | 1 |
| Query Match: | 65.51% | Indels: | 12 |
| DB: | 6 | Gaps: | 1 |

| | | | |
|---|-----|---|-----|
| US-09-938-703-3 (1-1442) x Q95NC7 (1-352) | | | |
| QY | 240 | ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTATATACATCGAGCCCTGC | 299 |
| DB | 1 | MetAspTyrGlnValSerSerProThrTyrAspIleAspTyrThrSerGluProCys | 20 |
| QY | 300 | CAAAATCAATGTGAAGCAAAATGCGACCGCGCTCCTCGCTCGCTACTACTGGTG | 359 |
| DB | 21 | GlnLysValAsnValLysGlnIleAlaAlaArgLeuLeuProLeuTyrSerLeuVal | 40 |
| QY | 360 | TTCATCTTTGGTTTGTGGGCAACATGCTGTCATCTCTCATCTCGATAAAGTCCAAAGG | 419 |
| DB | 41 | PheIlePheGlyPheValGlyAsnIleLeuValIleuIleLeuIleAsnCysLysArg | 60 |
| QY | 420 | CTGAAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTCT | 479 |
| DB | 61 | LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu | 80 |
| QY | 480 | CTTACTGTCCTCTCTGGCTCATTACTGCTGCGCCGACGTGGAGTCTTGAATCAATG | 539 |
| DB | 81 | LeuThrValProPheTrpAlaHisTyrAlaAlaArgTrpAspPheGlyAsnThrMet | 100 |
| QY | 540 | TGTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTGGAATCTCTTTCATCATC | 599 |
| DB | 101 | CysGlnLeuLeuThrGlyLeuTyrPheIleGlyPhePheSerGlyIlePhePheIleIle | 120 |
| QY | 600 | CTCCTGACAAATCGATAGGTACCTGCTGCTGCTCCATGCTGTGTGTTGTTTAAAGCCAGG | 659 |
| DB | 121 | LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg | 140 |
| QY | 660 | ACGGTCACCTTTGGGTGCTGACAGTCTGATGATCATTGGGTGCTGCTGTGTCGCTCT | 719 |
| DB | 141 | ThrValThrPheGlyValValThrSerValIleThrTrpValAlaValPheAlaSer | 160 |
| QY | 720 | CTCCCAGGAATCATCTTTACAGATCTCAAAAAGAGGTCTCATTTACACCTCGAGCTCT | 779 |
| DB | 161 | LeuProGlyIleIlePheThrArgSerGlnArgGlnGlyLeuHisTyrThrCysSerSer | 180 |
| QY | 780 | CATTTCCTAC-----AFTAAAGATAGTCAT | 806 |
| DB | 181 | HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValIle | 200 |
| QY | 807 | CTTGGGGCTGTGCTCGCTGCTGTGTCATGGTCACTGCTACTCGGGATCTCTAAAC | 866 |
| DB | 200 | eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysThr | 220 |

GN CCR5.
 OS Pygathrix avunculus (Tonkin snub-nosed monkey), and
 OS Pygathrix roxellana (golden snub-nosed monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Pygathrix.
 OX NCBI_TaxID=6062, 61622;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Y.-W., Zhang Y.-P.;
 RL "Sequence evolution of chemokine receptor CCR5 gene in primates";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF075447; AAD19859.1; -;
 DR EMBL; AF075444; AAD19856.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40482 MW; 037CFA9E12E532F3 CRC64;
 Alignment Scores:
 Pred. No.: 338e-161 Length: 352
 Score: 1717.50 Matches: 331
 Percent Similarity: 96.03% Conservative: 8
 Best Local Similarity: 93.77% Mismatches: 2
 Query Match: 65.35% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x 097962 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGATCAATCAATTAATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspPheGlyLeuValThrSerGluProCys 20
 QY 300 CAAAATAATCAATGTGAAGCAATGCGAGCCGCTCTGCTCGCTGCTACTACTGCTG 359
 Db 21 GlnLysValAsnValLysGlnLeuAlaArgLeuLeuProLeuTyrSerLeuVal 40
 QY 360 TTCATCTTGGTGTGGGCAACATGCTGTCATCCTCATCTGATCAATCAATGCAAGG 419
 Db 41 PheLeuPheGlyPheValGlyAsnLeuValValLeuLeuLeuLeuLeuLeuLeuLeu 60
 QY 420 CTGAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCT 479
 Db 61 LeuLysSerMetThrAspIleTyrLeuLeuAsnLeuAlaIleSerAspLeuPheLeu 80
 QY 480 CTTACTGTCCCTCTCTGGCTCCTATGCTGCTGCGCCGCTGCTGCTGCTGCTGCTG 539
 Db 81 LeuThrValProPheThrPalaHisTyrAlaAlaAlaGlnTyrPheGlyAsnThrMet 100
 QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTCTCTCATCATC 599
 Db 101 CysGlnLeuLeuThrGlyLeuTyrPheLeuGlyPheSerGlyLeuPhePheLeuLeu 120
 QY 600 CTCTCGACATCGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
 Db 121 LeuLeuThrIleAspArgTyrLeuAlaIleValHisAlaValPheAlaLeuLysAlaArg 140
 QY 660 AGGTCACCTTTGGGTGGTGACAAAGTGTGATCCTGCTGCTGCTGCTGCTGCTGCTG 719
 Db 141 ThrValThrPheGlyValValThrSerValIleThrTyrValValAlaValPheAlaSer 160
 QY 720 CTTCCAGGAATCACTTTACAGATCTCAAAAGAGGCTCTCATCTACCTGCTGCTGCT 779
 Db 161 LeuProGlyIleLeuPheThrArgSerGlnArgGluGlyValHisTyrThrCysSerSer 180
 QY 780 CATTTTCCATAC-----ATTAAAGTACTCAT 806
 Db 181 HisPheProTyrSerGlnTyrGlnPheThrPheGlnThrLeu-LysIleValIle 200
 QY 807 CTGGGGCTGGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866

Db 200 eLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
 QY 867 TCTGCTTCGCTCGAATCAG 926
 Db 220 rLeuLeuArgCysArgAsnGlnLysArgHisArgAlaValArgLeuLeuPheThrIle 240
 QY 927 CATGATTGTTTATTTCTCTCTGGGCTCCCTACACATTCCTCTCTCTCTCTCTCTCT 986
 Db 240 eMetIleValTyrPheLeuPheThrPalaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
 QY 987 CCAGGAATCTTGGCTGCTGAATTCAGTAGTCTACACAGTTGGACCAAGCTATGCA 1046
 Db 260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
 QY 1047 GGTGACAGACTCTTGGGATGACCTGCTGATCAACCCCATCATCTATGCTTCTCT 1106
 Db 280 nValThrGluThrLeuGlyMetThrHisCysIleAsnProIleIleTyrAlaPheVa 300
 QY 1107 CGGGAGAGAGTTCAGAACTACCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1166
 Db 300 lGlyGluLysPheArgAsnTyrLeuValPhePheGlnLysHisIleAlaLysArgPh 320
 QY 1167 CTGCAATCTGTTCTATTTTCCAGCAAGAGCTCCGAGGAGCAAGCTCAGTTTACAC 1226
 Db 320 eCysLysCysCysSerIlePheGlnGlnGlnAlaProGluArgAlaSerSerValTyrTh 340
 QY 1227 CCGATCCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263
 Db 340 rArgSerThrGlyGluGluGluThrSerValGlyLeu 352

RESULT 10
 Q9XTL4
 ID Q9XTL4 PRELIMINARY; PRT; 352 AA.
 AC Q9XTL4;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CC chemokine receptor 5.
 GN CCR5.
 OS Colobus guereza (Black-and-white colobus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
 OC Colobus.
 OX NCBI_TaxID=33548;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Spearman P.W., Mburu D.N., Graham B.S.;
 RL "Differential Utilization of CCR5 Molecules from Three East African
 RT Simian Species by the HIV-1 Envelope Glycoprotein".
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF141639; AAD32684.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 352 AA; 40550 MW; 9B078EF04D34DB36 CRC64;
 Alignment Scores:
 Pred. No.: 4.24e-161 Length: 352
 Score: 1716.50 Matches: 331
 Percent Similarity: 95.75% Conservative: 7
 Best Local Similarity: 93.77% Mismatches: 3
 Query Match: 65.32% Indels: 12
 DB: 6 Gaps: 1

US-09-938-703-3 (1-1442) x Q9XTL4 (1-352)

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGATCAATCAATTAATATACATCGAGCCCTGC 299
 Db 1 MetAspTyrGlnValSerProThrTyrAspPheGlyLeuValThrSerGluProCys 20


```

AC Q95NC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Miopithecus.
OX NCBI_TaxID=36231;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
  phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177886; AAK43369.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40546 MW; 6464152F3E566AE5 CRC64;

Alignment Scores:
Pred. No.: 5.33e-161 Length: 352
Score: 1715.50 Matches: 332
Percent Similarity: 95.75% Conservative: 6
Best Local Similarity: 94.05% Mismatches: 3
Query Match: 65.20% Indels: 12
DB: Gaps: 1

US-09-938-703-3 (1-1442) x Q95NC3 (1-352)

QY 240 ATGATTATCAAGTGTCAAGTCCAAATGATGATCAATATTATATACATCGGAGCCCTGC 299
DB 1 MetaspYrGlnValSerSerProThrYrAspIleAsnYrYrThrSerGluProCys 20
QY 300 CAAAATAATGTGAAGCAATCGAGCCGCCCTCCCTCGCTGCTACCTACCTGCTG 359
DB 21 GlnYrIleAsnValYrGlnIleAlaAlaArgLeuLeuProLeuYrSerLeuVal 40
QY 360 TTCATCTTGGTTTGTGGCAACATGCTGTCATCCATCCATCCATCAAGTCAAGG 419
DB 41 PheIlePheGlyPheValGlyAsnIleLeuValIleLeuValIleLeuIleAsnCysLysArg 60
QY 420 CTGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCT 479
DB 61 LeuYrSerMetThrAspIleYrLeuLeuAsnLeuAlaIleSerAspLeuLeuPheLeu 80
QY 480 CTTACTGTCCTTCGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
DB 81 LeuThrValProPheThrAlaHisYrAlaAlaAlaGlnTrpAspPheGlyAsnThrMet 100
QY 540 TGTCAACTCTGACAGGCTCTATTATATAGGCTTCTCTGGAATCTCTTCATCATC 599
DB 101 CysArgLeuLeuThrGlyLeuYrPheIleGlyPheSerGlyIlePhePheIleIle 120
QY 600 CTCCTGACAAATGATGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
DB 121 LeuLeuThrIleAspArgYrLeuAlaIleValHisAlaValPheAlaLeuYrAlaArg 140
QY 660 AGGTCACCTTGGGTTGTGACAGTGTGATCACTGCTGGTGGTGGTGGTGGTGGTGGTCT 719
DB 141 ThrValThrPheGlyValValThrSerValIleThrTrpValValAlaValPheAlaSer 160
QY 720 CTCACGAAATCATCTTACCATCTCAAGATCTCAAGATCTCAAGATCTCAAGATCTCA 779
DB 161 LeuProGlyIleIlePheThrArgSerGlnArgGlnGlyLeuHisYrThrCysSerSer 180
QY 780 CATTTTCCATAC-----ATTAAGATAGTCAAT 806
DB: ||||| |||||

181 HisPheProTyrSerGlnTyrGlnPheTrpLysAsnPheGlnThrLeu-LysIleValI 200
807 CTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
200 eLeuGlyLeuValLeuProLeuValMetValIleCysYrSerGlyIleLeuLysTh 220
867 TCTGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926
220 rLeuLeuArgCysArgAsnGlnLysLysArgHisArgAlaValArgLeuIlePheThrI 240
927 CATGATGCTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
240 eMetIleValYrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuLeuAsnThrPh 260
987 CCAGGAATCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
260 eGlnGluPhePheGlyLeuAsnAsnCysSerSerSerAsnArgLeuAspGlnAlaMetG 280
1047 GGTGACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleYrAlaPheVa 300
1107 CGGGGAGAAGTTCAGAACTACCTCTTAGTCTCTTCCAAAAGCACATTCGCCAAGCGCTT 1166
300 lGlyGluLysPheArgAsnYrLeuValPhePheGlnLysHisIleAlaLysArgPh 320
1167 CTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
320 eCysLysCysSerIlePheGlnGlnAlaProGluArgAlaSerSerValYrTh 340
1227 CGATCCACTGGGAGCAGCAATATCTGCTGGCTG 1263
340 rArgSerMetGlyGluGlnGluIleSerValGlyLeu 352

RESULT 14
Q9XT13
ID Q9XT13 PRELIMINARY; PRT; 352 AA.
AC Q9XT13;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CC chemokine receptor 5.
GN CCR5.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9553;
RN [1]
RP SEQUENCE FROM N.A.
RA Spearman P.W., Mburu D.N., Graham B.S.;
RT "Differential utilization of CCR5 molecules from Three East African
  Simian Species by the HIV-1 Envelope Glycoprotein.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141640; AAD32685.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; GPCR_Rhodpsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40489 MW; 0B47E337C11E2E1E CRC64;

Alignment Scores:
Pred. No.: 8.4e-161 Length: 352
Score: 1713.50 Matches: 332
Percent Similarity: 95.75% Conservative: 6
Best Local Similarity: 94.05% Mismatches: 3
Query Match: 65.20% Indels: 12
DB: Gaps: 1

US-09-938-703-3 (1-1442) x Q9XT13 (1-352)

```



```
Db      161 LeuProGlyllellePheThrArgSerGlnArgGluGlyLeuHisTyrThrCysSerSer 180
QY      780 CATTTCATAC-----ATTAAAGATAGTCAT 806
      |||||
Db      181 HisPheProTyrSerGlnTyrGlnPheTyrLysAsnPheGlnThrLeu-LysMetValI 200
      |||||
QY      807 CTTGGGCTGGTCTCGCTGCTGTCATGGTCATCTGCTACTCGGGAATCCTAAAC 866
      |||||
Db      200 eLeuGlyLeuValProLeuLeuValMetValIleCysTyrSerGlyIleLeuLysTh 220
      |||||
QY      867 TCTGCTTCGGTGTGGAATGAGAAGAGGACAGGCTGTGAGGCTTATCTCACCAT 926
      |||||
Db      220 rIeuLeuArgCysArgAsnGluLysArgHisArgAlaValArgLeuIlePheThrI 240
      |||||
QY      927 CAGATGTTTATTTCTCTCTGGCTCCCTACACATGCTCTCTCTCTGACACCTT 986
      |||||
Db      240 eMetIleValTyrPheLeuPheTrpAlaProTyrAsnIleValLeuLeuAsnThrPh 260
      |||||
QY      987 CCAGGAATCTTTGGCTCAATAATTGCAGTAGCTTAACAGGTGGACCAAGCTATGCA 1046
      |||||
Db      260 eGlnGluPhePheGlyLeuAsnAsnCysSerCysAsnArgLeuAspGlnAlaMetGl 280
      |||||
QY      1047 GGTGACAGAGACTCTTGGGATGACGCACTGCTGCATCAACCCATCATCTATGCTTTGT 1106
      |||||
Db      280 nValThrGluThrLeuGlyMetThrHisCysCysIleAsnProIleIleTyrAlaPheVa 300
      |||||
QY      1107 CCGGAGAGACTTCAGAACTACTCTTAGTCTTCTTCCAAAGCACATTGCCAAACGCTT 1166
      |||||
Db      300 lGlyGluLysPheArgAsnTyrLeuLeuValPhePheGlnLysHisIleAlaLysArgPh 320
      |||||
QY      1167 CTGCAATGCTGTCTATTTCACAGAGAGGCTCCGAGCGACGACAGCTCAGTTTACAC 1226
      |||||
Db      320 eCysLysCysSerIlePheGlnGlnGluAlaProGluArgAlaSerSerValTyrTh 340
      |||||
QY      1227 CCGATCCACTGGGAGCAGGAATATCTGTGGGCTTG 1263
      |||||
Db      340 rArgSerThrGlyGluGlnGluIleSerValGlyLeu 352
```

Search completed: June 3, 2003, 18:49:48
Job time : 119.268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: June 3, 2003, 18:39:24 ; Search time 4.86523 Seconds
(without alignments)
5149.005 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885
Perfect score: 164
Sequence: 1 ATTAAGATGATCACTTGG.....CTCTGCTCGGTGTCGAAT 94

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p.model -DB=xlp
-Q/cgn2_1/USPTO_pool/US0938703/runat_03062003_161406_22796/app_query.fasta_1.1870
-DB-A_Geneseq101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938703.qcgn_1.162 @runat_03062003_161406_22796 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NGB_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1 | 164 | 100.0 | 215 | 18 | AAW27408 |
| 2 | 164 | 100.0 | 215 | 20 | AAW88238 |
| 3 | 152 | 92.7 | 332 | 18 | AAW26766 |
| 4 | 152 | 92.7 | 352 | 18 | AAW27407 |
| 5 | 152 | 92.7 | 352 | 18 | AAW27123 |
| 6 | 152 | 92.7 | 352 | 18 | AAW07602 |
| 7 | 152 | 92.7 | 352 | 19 | AAW23835 |
| 8 | 152 | 92.7 | 352 | 20 | AAW88232 |
| 9 | 152 | 92.7 | 352 | 21 | AAW80128 |
| 10 | 152 | 92.7 | 352 | 22 | ABB56342 |
| 11 | 152 | 92.7 | 352 | 22 | AAW80111 |
| 12 | 152 | 92.7 | 352 | 22 | AAW82948 |
| 13 | 152 | 92.7 | 352 | 22 | AAW79089 |
| 14 | 152 | 92.7 | 352 | 22 | AAE07037 |
| 15 | 152 | 92.7 | 352 | 22 | AAE07039 |
| 16 | 152 | 92.7 | 352 | 22 | AAE07046 |
| 17 | 152 | 92.7 | 352 | 22 | AAE07048 |
| 18 | 152 | 92.7 | 352 | 22 | AAW83354 |
| 19 | 152 | 92.7 | 352 | 22 | AAE04321 |
| 20 | 152 | 92.7 | 352 | 22 | AAW46858 |
| 21 | 152 | 92.7 | 352 | 23 | AAW97150 |
| 22 | 152 | 92.7 | 352 | 23 | AAU97152 |
| 23 | 152 | 92.7 | 352 | 23 | ABB08343 |
| 24 | 152 | 92.7 | 352 | 23 | AAW52828 |
| 25 | 152 | 92.7 | 352 | 23 | AAW52829 |
| 26 | 152 | 92.7 | 371 | 19 | AAW23834 |
| 27 | 152 | 92.7 | 439 | 20 | AAW41280 |
| 28 | 149 | 90.9 | 352 | 18 | AAW27125 |
| 29 | 134 | 81.7 | 329 | 22 | AAW46859 |
| 30 | 134 | 81.7 | 360 | 16 | AAW79166 |
| 31 | 134 | 81.7 | 360 | 18 | AAW35833 |
| 32 | 134 | 81.7 | 360 | 22 | AAW56340 |
| 33 | 134 | 81.7 | 360 | 22 | AAW80108 |
| 34 | 134 | 81.7 | 360 | 22 | AAU07613 |
| 35 | 134 | 81.7 | 360 | 22 | AAU07614 |
| 36 | 134 | 81.7 | 374 | 16 | AAW79165 |
| 37 | 134 | 81.7 | 374 | 22 | AAW80107 |
| 38 | 131 | 79.9 | 354 | 19 | AAW54037 |
| 39 | 127 | 77.4 | 331 | 19 | AAW40027 |
| 40 | 109 | 66.5 | 332 | 23 | ABJ03698 |
| 41 | 109 | 66.5 | 355 | 17 | AAW03376 |
| 42 | 109 | 66.5 | 355 | 17 | AAW03377 |
| 43 | 109 | 66.5 | 355 | 17 | AAW03378 |
| 44 | 109 | 66.5 | 355 | 18 | AAW31850 |
| 45 | 109 | 66.5 | 355 | 18 | AAW27124 |

ALIGNMENTS

RESULT 1
AAW27408
ID AAW27408 standard; Protein; 215 AA.
XX
AC AAW27408;
XX
DT 14-APR-1998 (first entry)
XX
DE Inactive human CCR5.
XX
KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
OS Homo sapiens.
XX

Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215

QY 62 GGAATCCCAAAACTGCTTCGGTGCAGAAAT 94
 |||||
 Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 5

AAW27123
 ID AAW27123 standard; Protein; 352 AA.

XX AC AAW27123;

XX DT 14-DEC-1997 (first entry)

XX DE Human chemokine receptor 88C.

XX KW Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;

XX KW tumour; asthma; viral infection; AIDS; inflammation;

XX KW autoimmune disease; therapy; diagnosis; leukocyte trafficking;

XX KW G protein coupled receptor; ligand; modulator; antibody; human.

XX OS Homo sapiens.

XX EH Key Location/Qualifiers

XX FT Domain

XX FT 1..32

XX FT /label= Extracellular_domain

XX FT 56..67

XX FT /label= Intracellular_domain

XX FT 89..112

XX FT /label= Extracellular_domain

XX FT 125..145

XX FT /label= Intracellular_domain

XX FT 166..191

XX FT /label= Extracellular_domain

XX FT 213..235

XX FT /label= Intracellular_domain

XX FT 259..280

XX FT /label= Extracellular_domain

XX FT 301..352

XX FT /label= Intracellular_domain

XX PN WO9722698-A2.

XX PD 26-JUN-1997.

XX PF 20-DEC-1996; 96WO-US20759.

XX PR 07-JUN-1996; 96US-0661393.

XX PR 20-DEC-1995; 95US-0575967.

XX PA (ICOS-) ICOS CORP.

XX PI Gray PW, Raport CJ, Schweickart VL;

XX DR WPI; 1997-341689/31.

XX DR N-PSDB; AAT85161.

XX PT New nucleic acid encoding chemokine receptors 88-2B and 88C - used

XX PT to modulate leukocyte trafficking, e.g. for treatment of

XX PT inflammation, tumours, viral infections, autoimmune diseases, etc.

XX PS Claim 16; Page 47-48; 65pp; English.

XX CC This polypeptide sequence comprises novel human chemokine receptor

XX CC 88C, a G protein coupled receptor that is involved in leukocyte

XX CC trafficking. Its amino sequence was deduced from a cDNA clone

XX CC (AAT85161) isolated from a macrophage library. It shows 62% identity

XX CC to CCRKRI. Chemokine receptor 88-2B (see AAW27124) has also been

XX CC identified. 88C and 88-2B receptors and their polypeptide fragments

XX CC can be produced in transformed host cells. The receptors, peptides

XX CC comprising one or more of the extracellular or intracellular

XX CC domains, and anti-receptor antibodies can be used to modulate

XX CC receptor activities, particularly ligand and G protein binding, and

CC are potentially useful in the treatment of
 CC atherosclerosis, rheumatoid arthritis, tumours, asthma, viral
 CC infection, AIDS, inflammatory conditions, pathological immune
 CC response, abnormal haematopoietic processes etc.

XX SQ Sequence 352 AA;

Alignment Scores:

Pred. No.: 4.4e-13 Length: 352
 Score: 152.00 Matches: 31
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.68% Indels: 0
 DB: 18 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x AAW27123 (1-352)

QY 2 TTAAGATAGTCACTTGGGGTGGTCTGCGTCTGCTGCTCATGCTACTGCTACTG 61

Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215

QY 62 GGAATCCTAAACTCTGCTTCGGTGCAGAAAT 94

Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 6

AAW07602

ID AAW07602 standard; Protein; 352 AA.

XX AC AAW07602;

XX DT 26-FEB-1997 (first entry)

XX DE Human G-protein chemokine receptor HDGNR10.

XX KW G-protein chemokine receptor; HDGNR10; signal transduction;

XX KW haematopoiesis; leukaemia; inflammation; rheumatoid arthritis;

XX KW diagnosis; therapy.

XX OS Homo sapiens.

XX PN WO9639437-A1.

XX PD 12-DEC-1996.

XX PF 06-JUN-1995; 95WO-US07173.

XX PR 06-JUN-1995; 95WO-US07173.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Li Y, Ruben SM;

XX DR WPI; 1997-043072/04.

XX DR N-PSDB; AAT44042.

XX PT Human G-protein chemokine receptor, HDGNR10 - useful to identify

XX PT (antagonists, for treatment of haematopoiesis, leukaemia, chronic

XX PT inflammation, rheumatoid arthritis, etc.

XX PS Claim 1; Page 44-46; 61pp; English.

XX CC Novel human mature G-protein chemokine receptor HDGNR10 (AAW07602)

XX CC is a 7-transmembrane protein involved in signal transduction. Its

XX CC amino acid sequence was deduced from a cDNA clone (AAT44042) isolated

XX CC from a human monocyte library. Isolation of the cDNA allows

XX CC prodn. of recombinant HDGNR10 in host, e.g. E. coli, COS or Sf9,

XX CC cells. The recombinant HDGNR10 can be used to identify agonists

XX CC or antagonists of the receptor; such cpds. can be used to treat

XX CC conditions related to the under- and over-expression of G-protein

XX CC chemokine receptors.

XX SQ Sequence 352 AA;

RESULT 11
AAG80111
Db 216 GlyLeuLeuLysThrLeuLeuArgCysArgAsn 216
02 GGATTCCTAATAAAACCTCGCTTCGGTGCAGAAAT 94

QY 62 GGATCTCTAAACCTCTGCTCGTCTCGTCTCGAAT 94
|||||
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
RESULT 14
AAE07037
ID AAE07037 standard; Protein; 352 AA.
AC AAE07037;
XX
XX 16-OCT-2001 (first entry)
XX Human G-protein chemokine receptor (CCR5) HDGNR10 protein #1.
XX Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;
KW human immunodeficiency virus; antimicrobial; vasodilator; vulnery;
KW cytostatic; immunosuppressive; nootropic; neuroprotective; gene therapy;
KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;
KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;
KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;
KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;
KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;
KW cardiovascular disorder; myocardial ischaemia.
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH Domain 1..36
FT /label= Extracellular_domain
FT Domain 37..305
FT /label= Transmembrane_domain
FT Domain 37..58
FT /label= Transmembrane_domain
FT /note= "Segment 1"
FT Domain 59..67
FT /label= Intracellular_loop_1
FT Domain 68..88
FT /label= Transmembrane_domain
FT /note= "Segment 2"
FT Domain 89..102
FT /label= Extracellular_loop_1
FT Domain 103..124
FT /label= Transmembrane_domain
FT /note= "Segment 3"
FT Domain 125..141
FT /label= Intracellular_loop_2
FT Domain 142..166
FT /label= Transmembrane_domain
FT /note= "Segment 4"
FT Domain 167..195
FT /label= Extracellular_loop_2
FT Domain 196..223
FT /label= Transmembrane_domain
FT /note= "Segment 5"
FT Domain 224..235
FT /label= Intracellular_loop_3
FT Domain 236..260
FT /label= Transmembrane_domain
FT /note= "Segment 6"
FT Domain 261..274
FT /label= Extracellular_loop_3
FT Domain 287..305
FT /label= Transmembrane_domain
FT /note= "Segment 7"
FT Domain 306..352
FT /label= Intracellular_domain
XX
XX WO200158915-A2.
XX
XX 16-AUG-2001.
XX
XX 09-FEB-2001; 2001WO-US04152.
XX

PR 09-FEB-2000; 2000US-0181258.
PR 09-MAR-2000; 2000US-0187999.
PR 22-SEP-2000; 2000US-0234336.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Roschke V, Li Y, Ruben SM;
PI WPI; 2001-488965/53.
XX N-PSDB; AAD13181.
XX
XX Isolated nucleic acid encoding a human G-protein chemokine receptor
PT (CCR5) HDGNR10 polypeptide, useful for preventing or treating
PT autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
PT disorders and neurodegenerative disorders -
XX
XX Claim 102; Fig 1; 495pp; English.
XX
XX The present sequence is human G-protein chemokine receptor (CCR5) HDGNR10
CC protein. CCR5 HDGNR10 cDNA is included in ATCC Deposit No: 97183. CCR5
CC HDGNR10 antibodies are useful for treating, preventing or ameliorating a
CC disease or disorder associated with inflammation, defective or aberrant
CC chemotaxis of immune cells, HIV infection (such as Pneumocystis carinii
CC pneumonia or Kaposi's sarcoma) or defective or aberrant T-cell antigen
CC presenting cell interaction. The disease or disorder may also be an
CC infectious disease (e.g. a viral infection such as an early stage HIV
CC infection, a cytomegalovirus infection, or a poxvirus infection), an
CC autoimmune disease (e.g. rheumatoid arthritis) or a neurodegenerative
CC disorder. The disease or disorder may be associated with aberrant CCR5
CC expression, lack of CCR5 function, aberrant CCR5 ligand expression, or
CC lack of CCR5 ligand function. CCR5 HDGNR10 protein is used as a food
CC additive or preservative to increase or decrease storage capabilities.
CC CCR5 HDGNR10 DNA are useful for chromosome identification and in gene
CC therapy. CCR5 HDGNR10 DNA, protein, antibodies, agonists and antagonists
CC are also useful in the diagnosis, treatment and prevention of cancer
CC (breast, ovary, adrenal gland, bone, bone marrow, gastrointestinal tract,
CC liver, lung, urogenital); immune disorders (Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis); cardiovascular disorders (myocardial ischaemia) and
CC wound healing.
XX
XX SQ Sequence 352 AA;
Alignment Scores:
Pred. No.: 4.4e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 22 Gaps: 0
US-09-938-703-3_COPY_792_885 (1-94) x AAE07037 (1-352)
QY 2 TTAAGATAGTCACTTGGGGTGTCTCGCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 61
|||||
Db 196 LeuLysIleValIleLeuLysValLeuProLeuValMetValIleCysIleCysIle 215
QY 62 GGATCTCTAAACCTCTGCTCGTCTCGTCTCGAAT 94
|||||
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
RESULT 15
AAE07039
ID AAE07039 standard; Protein; 352 AA.
XX
XX AC AAE07039;
XX
XX 16-OCT-2001 (first entry)
XX
XX Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.
XX
XX Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:44:09 : Search time 1.80534 Seconds
(without alignments)
3063.973 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164
Sequence: 1 ATTAAGATAGTACTCTGG.....CTCGTCTGGTGTGCGAAT 94

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p_model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US0938703/runat_03062003_161408_22885/app_query.fasta_1.1870
-DB=Issued Patents AA -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0938703.ecgn_1_1_42 -runat_03062003_161408_22885 -NCP0=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|-------------------|-------------------|
| 1 | 164 | 100.0 | 34 | US-08-833-752-13 | Sequence 13, Appl |
| 2 | 164 | 100.0 | 215 | US-09-087-232A-17 | Sequence 17, Appl |
| 3 | 164 | 100.0 | 215 | US-08-833-752-6 | Sequence 6, Appl |
| 4 | 152 | 92.7 | 54 | US-08-833-752-11 | Sequence 11, Appl |
| 5 | 152 | 92.7 | 352 | US-08-466-343D-2 | Sequence 2, Appl |
| 6 | 152 | 92.7 | 352 | US-09-087-232A-13 | Sequence 13, Appl |
| 7 | 152 | 92.7 | 352 | US-08-861-105-14 | Sequence 14, Appl |
| 8 | 152 | 92.7 | 352 | US-08-575-967A-2 | Sequence 2, Appl |
| 9 | 152 | 92.7 | 352 | US-09-045-583-52 | Sequence 52, Appl |
| 10 | 152 | 92.7 | 352 | US-09-517-605-5 | Sequence 5, Appl |
| 11 | 152 | 92.7 | 352 | US-09-534-185-52 | Sequence 52, Appl |
| 12 | 152 | 92.7 | 352 | US-08-833-752-5 | Sequence 5, Appl |

| | | | | | | |
|----|-----|------|-----|---|-------------------|-------------------|
| 13 | 134 | 81.7 | 344 | 3 | US-08-466-343D-9 | Sequence 9, Appl |
| 14 | 134 | 81.7 | 347 | 1 | US-08-461-244-3 | Sequence 3, Appl |
| 15 | 134 | 81.7 | 360 | 1 | US-08-450-393A-4 | Sequence 4, Appl |
| 16 | 134 | 81.7 | 360 | 4 | US-08-446-669-4 | Sequence 4, Appl |
| 17 | 134 | 81.7 | 360 | 4 | US-09-045-583-50 | Sequence 50, Appl |
| 18 | 134 | 81.7 | 360 | 4 | US-09-045-583-51 | Sequence 51, Appl |
| 19 | 134 | 81.7 | 360 | 4 | US-09-534-185-50 | Sequence 50, Appl |
| 20 | 134 | 81.7 | 360 | 4 | US-08-833-752-7 | Sequence 7, Appl |
| 21 | 134 | 81.7 | 360 | 5 | PCT-US95-00476-4 | Sequence 4, Appl |
| 22 | 134 | 81.7 | 360 | 5 | PCT-US95-00476-4 | Sequence 2, Appl |
| 23 | 134 | 81.7 | 374 | 4 | US-08-446-669-2 | Sequence 2, Appl |
| 24 | 134 | 81.7 | 374 | 5 | PCT-US95-00476-2 | Sequence 2, Appl |
| 25 | 134 | 81.7 | 374 | 5 | PCT-US95-00476-2 | Sequence 2, Appl |
| 26 | 131 | 79.9 | 354 | 4 | US-08-724-984A-2 | Sequence 21, Appl |
| 27 | 115 | 70.1 | 24 | 4 | US-09-087-232A-21 | Sequence 4, Appl |
| 28 | 109 | 66.5 | 355 | 4 | US-08-575-967A-4 | Sequence 1, Appl |
| 29 | 109 | 66.5 | 355 | 4 | US-08-847-296B-1 | Sequence 54, Appl |
| 30 | 109 | 66.5 | 355 | 4 | US-09-045-583-54 | Sequence 54, Appl |
| 31 | 109 | 66.5 | 355 | 4 | US-09-534-185-54 | Sequence 54, Appl |
| 32 | 107 | 65.2 | 353 | 4 | US-08-833-752-8 | Sequence 8, Appl |
| 33 | 106 | 64.6 | 355 | 1 | US-08-012-988A-2 | Sequence 2, Appl |
| 34 | 106 | 64.6 | 355 | 1 | US-08-450-393A-5 | Sequence 5, Appl |
| 35 | 106 | 64.6 | 355 | 4 | US-08-446-669-5 | Sequence 5, Appl |
| 36 | 106 | 64.6 | 355 | 4 | US-09-239-938-1 | Sequence 1, Appl |
| 37 | 106 | 64.6 | 355 | 5 | PCT-US95-00476-5 | Sequence 5, Appl |
| 38 | 105 | 64.0 | 355 | 4 | US-09-045-583-53 | Sequence 53, Appl |
| 39 | 105 | 64.0 | 355 | 4 | US-09-534-185-53 | Sequence 53, Appl |
| 40 | 103 | 62.8 | 355 | 4 | US-08-833-752-9 | Sequence 9, Appl |
| 41 | 95 | 57.9 | 360 | 4 | US-08-875-573-20 | Sequence 20, Appl |
| 42 | 95 | 57.9 | 360 | 4 | US-09-232-878-2 | Sequence 2, Appl |
| 43 | 95 | 57.9 | 360 | 4 | US-09-045-583-55 | Sequence 55, Appl |
| 44 | 95 | 57.9 | 360 | 4 | US-09-534-185-55 | Sequence 55, Appl |
| 45 | 95 | 57.9 | 360 | 4 | US-08-833-752-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1

US-08-833-752-13

; Sequence 13, Application US/08833752

; Patent No. 6448375

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; APPLICANT: PARMENTIER, MARC

; APPLICANT: VASSART, GILBERT

; APPLICANT: LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/833,752

; FILING DATE: 9-APR-1997

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

; REFERENCE/DOCKET NUMBER:

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 34 amino acids

; TYPE: amino acid

US-08-833-752-11
; Sequence 11, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:

```
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-833-752-11

Alignment Scores:
Pred. No.: 3,13e-14 Length: 54
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-833-752-11 (1-54)
QY 2 TTAAGATAGTCATCTGGGGCTGCTCGCGCTGCTGTCATGTCATGCTACTGCG 61
Db 20 LeuYsIleValIleLeuGlyLeuValProLeuLeuValMetValIleCysTyrSer 39
QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
Db 40 GlyIleLeuYsThrLeuLeuArgCysArgAsn 50

RESULT 5
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-343D-2

Alignment Scores:
Pred. No.: 3,52e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 3 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-466-343D-2 (1-352)
QY 2 TTAAGATAGTCATCTGGGGCTGCTCGCGCTGCTGTCATGTCATGCTACTGCG 61
Db 196 LeuYsIleValIleLeuGlyLeuValProLeuLeuValMetValIleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
Db 216 GlyIleLeuYsThrLeuLeuArgCysArgAsn 226

RESULT 6
US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REGISTRATION NUMBER: 35,225
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
```


Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Matches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-575-967A-2 (1-352)

QY 2 TTAAGATAGTCATCTGGGCGTGGTCTGCTGCTCATGTCATGCTACTCG 61
Db 196 LeuylsileValleLeuGlyLeuValleProLeuValMetValleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
Db 216 GlylleLeuLythrLeuLeuArgCysArgasn 226

RESULT 9

US-09-045-583-52
; Sequence 52, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-045-583-52

Alignment Scores:
Pred. No.: 352e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Matches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-045-583-52 (1-352)

QY 2 TTAAGATAGTCATCTGGGCGTGGTCTGCTGCTCATGTCATGCTACTCG 61
Db 196 LeuylsileValleLeuGlyLeuValleProLeuValMetValleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94

Db 216 GlylleLeuLythrLeuLeuArgCysArgasn 226

RESULT 10

US-09-517-605-5
; Sequence 5, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtendeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-5

Alignment Scores:
Pred. No.: 3.52e-14 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 4 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-517-605-5 (1-352)

QY 2 TTAAGATAGTCATCTGGGCGTGGTCTGCTGCTCATGTCATGCTACTCG 61
Db 196 LeuylsileValleLeuGlyLeuValleProLeuValMetValleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGGTGCGAAAT 94
Db 216 GlylleLeuLythrLeuLeuArgCysArgasn 226

RESULT 11

US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptahelical Receptor Superfamily and Uses
; Therefor
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/534,185
; FILING DATE: 24-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,583
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

Alignment Scores:
Pred. No.:      3.52e-14      Length:      352
Score:          152.00       Matches:     31
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:      92.68%     Indels:        0
DB:               4         Gaps:          0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-833-752-5 (1-352)

QY      2  TTAAGATAGTCATCTTGGGCGTGCTCCGCGCTGTGTCATGTGCATCCTACTCG 161
Db      196 LeuLysIleValIleLeuGlyLeuValProLeuLeuValMetValIleCysTyrSer 215

QY      62 GGAACTCTAAAAAAGCTGCTTCGGTGTGCAAT 94
Db      216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 13
US-08-466-343D-9
; Sequence 9, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, Yi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; TITLE OF INVENTION: CHEKOKINE RECEPTOR HDGNRI0 (AS AMENDED)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466.343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-466-343D-9

Alignment Scores:
Pred. No.:      1.41e-11      Length:      344
Score:          134.00       Matches:     26
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 96.30% Mismatches:    0
Query Match:      81.71%     Indels:        0
DB:               3         Gaps:          0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-466-343D-9 (1-344)

QY      14  ATCTGGGCGTGCCTCCGCGCTGTTGTCATGTCATCTACTCGGAATCTCTAAA 73
Db      191 IleLeuGlyIleValLeuProLeuLeuIleMetValIleCysTyrSerGlyIleLeu 215

```

QY 74 ACTCTGCTCGTGTGCAAT 94
|||||
Db 211 ThrLeuLeuArgCysArgAsn 217

RESULT 14
US-08-461-244-3
; Sequence 3, Application US/08461244
; Patent No. 5776729
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Yi, Li
; APPLICANT: Ruben, Steven M.
; APPLICANT: Rosen, Craig A.

; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HGBR32
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,244
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-445
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-244-3

Alignment Scores:
Pred. No.: 1,41e-11 Length: 347
Score: 134.00 Matches: 26
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 96.30% Mismatches: 0
Query Match: 81.71% Indels: 0
DB: 1 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-461-244-3 (1-347)

QY 14 ACTCTGGGCTGCTCGCGCTGTGTCATGTCATCTGCTACTCGGAATCTCTAAAA 73
|||||
Db 195 ThrLeuGlyLeuValLeuProLeuLeuMetValLeuCysTyrSerGlyLeuLeuLys 214

QY 74 ACTCTGCTCGTGTGCAAT 94
|||||
Db 215 ThrLeuLeuArgCysArgAsn 221

RESULT 15
US-08-450-393A-4
; Sequence 4, Application US/08450393A
; Patent No. 5707815
; GENERAL INFORMATION:
; APPLICANT: Charo, Israel

; APPLICANT: Coughlin, Shaun
; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
; TITLE OF INVENTION: PROTEIN RECEPTORS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,393A
; FILING DATE: May 25, 1995
; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: UCAL-237/02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5165
; TELEFAX: 415-8857-0663

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-450-393A-4

Alignment Scores:
Pred. No.: 1,42e-11 Length: 360
Score: 134.00 Matches: 26
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 96.30% Mismatches: 0
Query Match: 81.71% Indels: 0
DB: 1 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-08-450-393A-4 (1-360)

QY 14 ACTTTGGGGCTGTCCTCGCGCTGTGTCATGTCATCTGCTACTCGGAATCTCTAAAA 73
|||||
Db 208 ThrLeuGlyLeuValLeuProLeuLeuMetValLeuCysTyrSerGlyLeuLeuLys 227

QY 74 ACTCTGCTCGTGTGCAAT 94
|||||
Db 228 ThrLeuLeuArgCysArgAsn 234

Search completed: June 3, 2003, 18:52:48

Job time : 2.80534 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:21:31 ; Search time 31.4634 Seconds
(without alignments)
201.057 Million cell updates/sec

Title: US-09-938-703-6
Perfect score: 1122
Sequence: 1 MDYQSSPIYDINITYTSEPC.....AACHGHLNPKNSASVSK 215

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_5/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_5/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_5/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_5/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_5/ptodata/1/iaa/PCUTUS_COMB.pep.*
6: /cgn2_5/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | | | | |
|------------|-------|-------------|--------|----|-------------------|-------------------|--|--|--|--|
| Result No. | Score | Query Match | Length | DB | ID | Description | | | | |
| | | | | | | | | | | |
| 1 | 1122 | 100.0 | 215 | 4 | US-09-087-232A-17 | Sequence 17, Appl | | | | |
| 2 | 1122 | 100.0 | 215 | 4 | US-08-833-752-6 | Sequence 6, Appl | | | | |
| 3 | 958 | 85.4 | 184 | 4 | US-08-833-752-4 | Sequence 4, Appl | | | | |
| 4 | 958 | 85.4 | 352 | 4 | US-09-087-232A-13 | Sequence 13, Appl | | | | |
| 5 | 958 | 85.4 | 352 | 4 | US-08-831-105-14 | Sequence 14, Appl | | | | |
| 6 | 958 | 85.4 | 352 | 4 | US-08-575-967A-2 | Sequence 2, Appl | | | | |
| 7 | 958 | 85.4 | 352 | 4 | US-08-833-752-5 | Sequence 5, Appl | | | | |
| 8 | 952 | 84.8 | 352 | 4 | US-09-045-583-52 | Sequence 52, Appl | | | | |
| 9 | 952 | 84.8 | 352 | 4 | US-09-534-185-52 | Sequence 52, Appl | | | | |
| 10 | 943 | 84.0 | 352 | 3 | US-08-466-343D-2 | Sequence 2, Appl | | | | |
| 11 | 936 | 83.4 | 352 | 4 | US-09-517-605-5 | Sequence 5, Appl | | | | |
| 12 | 776 | 69.2 | 354 | 4 | US-08-724-984A-2 | Sequence 2, Appl | | | | |
| 13 | 694 | 61.9 | 344 | 3 | US-08-466-343D-9 | Sequence 9, Appl | | | | |
| 14 | 694 | 61.9 | 347 | 1 | US-08-461-244-3 | Sequence 3, Appl | | | | |
| 15 | 694 | 61.9 | 360 | 1 | US-08-450-393A-4 | Sequence 4, Appl | | | | |
| 16 | 694 | 61.9 | 360 | 4 | US-08-446-669-4 | Sequence 4, Appl | | | | |
| 17 | 694 | 61.9 | 360 | 4 | US-09-045-583-50 | Sequence 50, Appl | | | | |
| 18 | 694 | 61.9 | 360 | 4 | US-09-534-185-50 | Sequence 50, Appl | | | | |
| 19 | 694 | 61.9 | 360 | 5 | PCN-US95-00476-4 | Sequence 4, Appl | | | | |
| 20 | 694 | 61.9 | 374 | 1 | US-08-450-393A-2 | Sequence 2, Appl | | | | |
| 21 | 694 | 61.9 | 374 | 4 | US-08-446-669-2 | Sequence 2, Appl | | | | |
| 22 | 694 | 61.9 | 374 | 5 | PCN-US95-00476-2 | Sequence 2, Appl | | | | |
| 23 | 692 | 61.7 | 360 | 4 | US-08-833-752-7 | Sequence 7, Appl | | | | |
| 24 | 688 | 61.3 | 360 | 4 | US-09-045-583-51 | Sequence 51, Appl | | | | |
| 25 | 688 | 61.3 | 360 | 4 | US-09-534-185-51 | Sequence 51, Appl | | | | |
| 26 | 598 | 53.3 | 355 | 1 | US-08-012-988A-2 | Sequence 2, Appl | | | | |
| 27 | 598 | 53.3 | 355 | 1 | US-08-450-393A-5 | Sequence 5, Appl | | | | |

28 598 53.3 355 4 US-08-446-669-5 Sequence 5, Appli
29 598 53.3 355 4 US-09-239-938-1 Sequence 1, Appli
30 598 53.3 355 5 PCT-US95-00476-5 Sequence 5, Appli
31 594 52.9 355 4 US-08-833-752-9 Sequence 9, Appli
32 562 50.1 355 4 US-09-045-583-53 Sequence 53, Appli
33 562 50.1 355 4 US-09-534-185-53 Sequence 53, Appli
34 547.5 48.8 360 4 US-08-875-573-20 Sequence 20, Appli
35 547.5 48.8 360 4 US-09-232-878-2 Sequence 2, Appli
36 547.5 48.8 360 4 US-09-045-583-55 Sequence 55, Appli
37 547.5 48.8 360 4 US-09-534-185-55 Sequence 55, Appli
38 525 46.8 100 4 US-09-087-232A-15 Sequence 15, Appli
39 515.5 45.9 360 4 US-08-833-752-10 Sequence 10, Appli
40 506.5 45.1 355 4 US-08-875-967A-4 Sequence 4, Appli
41 506.5 45.1 355 4 US-08-847-296B-1 Sequence 1, Appli
42 506.5 45.1 355 4 US-09-045-583-54 Sequence 54, Appli
43 506.5 45.1 355 4 US-09-534-185-54 Sequence 54, Appli
44 502.5 44.8 355 4 US-08-833-752-8 Sequence 8, Appli
45 448 39.9 87 4 US-09-087-232A-18 Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-09-087-232A-17
; Sequence 17, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,232A
; FILING DATE: 28 MAY 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,057
; FILING DATE: 30 MAY 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: KOLE, LISA B.
; REFERENCE/DOCKET NUMBER: AP 31115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 408-2628
; TELEFAX: (212) 765-2519
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-087-232A-17

Query Match 100.0%; Score 1122; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.4e-102;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVFIFFGVGNMVLILNCKR 60
|||||
Db 1 MDYQSSPIYDINITYTSEPCQKINVKQIAARLLPPLYSLVFIFFGVGNMVLILNCKR 60
|||||

QY 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
QY 121 LITDRIYLVAVHAFKARTVTEGVTSTVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LITDRIYLVAVHAFKARTVTEGVTSTVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASYSK 215
Db 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASYSK 215

RESULT 2

US-08-833-752-6
; Sequence 6, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-6

Query Match 100.0%; Score 1122; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.4e-102;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
QY 121 LITDRIYLVAVHAFKARTVTEGVTSTVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LITDRIYLVAVHAFKARTVTEGVTSTVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASYSK 215
Db 181 HPPIKSHLGAGPAAACHGHLILGNPKNSASYSK 215

RESULT 3

US-08-833-752-4
; Sequence 4, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 184 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-833-752-4

Query Match 85.4%; Score 958; DB 4; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.9e-86;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
Db 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLTVPFWAHYAAQWDFGNTMCQLLTGLYFGFFSGIFII 120
QY 121 LITDRIYLVAVHAFKARTVTEGVTSTVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
Db 121 LITDRIYLVAVHAFKARTVTEGVTSTVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
QY 181 HPPI 184
Db 181 HPPI 184

RESULT 4

US-09-087-232A-13
; Sequence 13, Application US/09087232A
; Patent No. 6153431
; GENERAL INFORMATION:
; APPLICANT: Quillent et al.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
; TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

RESULT 5
 US-08-861-105-14
 ; Sequence 14, Application US/08861105
 ; Patent No. 6258527
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, RONG
 ; APPLICANT: DENG, HONGKUI
 ; APPLICANT: ELMETER, WILFRIED
 ; APPLICANT: LANDAU, NATHANIEL R.
 ; APPLICANT: LIU, RONG
 ; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
 ; TITLE OF INVENTION: MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
 ; TITLE OF INVENTION: USES THEREOF
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; STREET: Floor
 ; CITY: Hackensack

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-861-105-14

Query Match      85.4%; Score 958; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-85;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MDXVSSPIYDINNYISEPCOKINVKQIARLLPPLSLVFIIEGFGVMVLILLINCKR 60
Db      1 MDXVSSPIYDINNYISEPCOKINVKQIARLLPPLSLVFIIEGFGVMVLILLINCKR 60

QY      61 LKSWTDIYLLMAISDLFFLLTVFPFWAHYAAQADFGNTWCQLLTGLTFIGFTSGIFIT 120
Db      61 LKSWTDIYLLMAISDLFFLLTVFPFWAHYAAQADFGNTWCQLLTGLTFIGFTSGIFIT 120

QY      121 LLRIDRYLVAVHAFVFAKARTVTGGVTSVTIYVYVAFSLPIGIFTRSQEGLHYTCSS 180
Db      121 LLRIDRYLVAVHAFVFAKARTVTGGVTSVTIYVYVAFSLPIGIFTRSQEGLHYTCSS 180

```

ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/575,967A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6265184and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32918
TELEPHONE: 206-485-1900
TELEFAX: 206-485-1662
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: /- "88C amino acid sequence"
US-08-575-967A-2

Query Match 85.4%; Score 958; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-85;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMLVILLINCKR 60
|||||
Db 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
|||||
Db 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
QY 121 LLTIDRYLAVHVAHFALKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
|||||
Db 121 LLTIDRYLAVHVAHFALKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
QY 181 HPFY 184
|||||
Db 181 HPFY 184

RESULT 7
US-08-833-752-5
Sequence 5, Application US/08833752
Patent No. 6448375
GENERAL INFORMATION:
APPLICANT: SAMSON, MICHEL
APPLICANT: PARMENTIER, MARC
APPLICANT: VASSART, GILBERT
APPLICANT: LIBERT, FREDERICK
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,752
FILING DATE: 9-APR-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-752-5
Query Match 85.4%; Score 958; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-85;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMLVILLINCKR 60
|||||
Db 1 MDYQSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSILVFIFGVGNMLVILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
|||||
Db 61 LKSMTDIYLLNLAIISDLFFLLTPFPWAHYAAQWDFGNTMCQLTGLYFGFFSGIFFII 120
QY 121 LLTIDRYLAVHVAHFALKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
|||||
Db 121 LLTIDRYLAVHVAHFALKARTVFGVTVSVITWVAVFASLPGLIIFTRSKQGLHYTCSS 180
QY 181 HPFY 184
|||||
Db 181 HPFY 184
RESULT 8
US-09-045-583-52
Sequence 52, Application US/09045583
Patent No. 6287805
GENERAL INFORMATION:
APPLICANT: Graham, Gerard J. et al.
TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,583
FILING DATE: 20-MAR-98
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: ANI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:

;
; LENGTH: 352 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-045-583-52

Query Match 84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSVIFGVGNMVLILLINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSVIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCQLLTGLYFGFFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCQLLTGLYFGFFSGIFFII 120
QY 121 LTIIDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
DB 121 LTIIDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 9
US-09-534-185-52
; Sequence 52, Application US/09534185
; Patent No. 6403767
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6403767el Molecules of the G Protein-Coupled
; Heptahelical Receptor Superfamily and Uses
; Therefor

NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/534,185
FILING DATE: 24-Mar-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/045,583
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-044
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-534-185-52

Query Match 84.8%; Score 952; DB 4; Length 352;
Best Local Similarity 98.9%; Pred. No. 3.9e-85;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSVIFGVGNMVLILLINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSVIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCQLLTGLYFGFFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCQLLTGLYFGFFSGIFFII 120
QY 121 LTIIDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
DB 121 LTIIDRYLAVVHAFKARTVTFGVVTSVITWVAVFASLPGIIFTRSQKGLHYTCSS 180
QY 181 HFPY 184
DB 181 HFPY 184

RESULT 10
US-08-466-343D-2
; Sequence 2, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,343D
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 35,688
REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-343D-2

Query Match 84.08%; Score 943; DB 3; Length 352;
Best Local Similarity 97.8%; Pred. No. 2.9e-84;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSVIFGVGNMVLILLINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLSVIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCQLLTGLYFGFFSGIFFII 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPFWAHYAAAQWDFGNTMCQLLTGLYFGFFSGIFFII 120

QY 121 LITDRYLAVHAFKARTVTFGVTSVITWVAVFASLPGIIFRSQKGLHYTCSS 180
Db 121 LITDRYLAVHAFKARTVTFGVTSVITWVAVFASLPGIIFRSQKGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 11
US-09-517-605-5
; Sequence 5, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Tneo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-5

Query Match 83.4%; Score 936; DB 4; Length 352;
Best Local Similarity 97.3%; Pred. No. 1.4e-83;
Matches 179; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDYQVSSPIYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLIVLINCKR 60
Db 1 MDYQVSSPTDIDYDSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLIVLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFTII 120
Db 61 LKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFTII 120
QY 121 LITDRYLAVHAFKARTVTFGVTSVITWVAVFASLPGIIFRSQKGLHYTCSS 180
Db 121 LITDRYLAVHAFKARTVTFGVTSVITWVAVFASLPGIIFRSQKGLHYTCSS 180
QY 181 HFPY 184
Db 181 HFPY 184

RESULT 12
US-08-724-984A-2
; Sequence 2, Application US/08724984A
; Patent No. 6388055
; GENERAL INFORMATION:
; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-
; TITLE OF INVENTION: CKR5 Receptor
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road, P.O. Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

SOFTWARE: MICROSOFT WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,984A
FILING DATE: October 3, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: William T. Han
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50023
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270 5024
TELEFAX: 610 270 5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 354
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-724-984A-2
Query Match 69.2%; Score 776; DB 4; Length 354;
Best Local Similarity 79.0%; Pred. No. 5.3e-68;
Matches 147; Conservative 17; Mismatches 20; Indels 2; Gaps 1;
QY 1 MDYQVSSP--IYDINTYSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLIVLINCK 58
Db 1 MDYQVSSPTDIDYDGNMSPCQKINVKQIAARLLPPLYSLVIFGFGAGNMVFLILISC 60
QY 59 KRLKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFT 118
Db 61 KRLKSMTDIYLLNLAIISDLFFLLVFPWAHYAAQWDFGNTMCQLLTGLYFGFGSIFFT 120
QY 119 ILLITDRYLAVHAFKARTVTFGVTSVITWVAVFASLPGIIFRSQKGLHYTC 178
Db 121 ILLITDRYLAVHAFKARTVTFGVTSVITWVAVFASLPGIIFRSQKGLHYTC 180
QY 179 SSHPY 184
Db 181 SPHPPH 186
RESULT 13
US-08-466-343D-9
; Sequence 9, Application US/08466343D
; Patent No. 6025154
; GENERAL INFORMATION:
; APPLICANT: LI, Yi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN
; TITLE OF INVENTION: CHEMOKINE RECEPTOR HDGNR10 (AS AMENDED)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,343D
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM
; TELECOMMUNICATION INFORMATION:

QUALITY MATCH
01.9%; SCORE 094; DB I; Length 34/;

142 IVHAVFALKARTVTFGVVTSVITWLVAVFASVPGIIFTKCQKEDSVYVCGPYFP 195

Search completed: June 3, 2003, 19:25:10
Job time : 32.4634 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:17:26 ; Search time 55.061 Seconds
(without alignments)
520.312 Million cell updates/sec

Title: US-09-938-703-6

Perfect score: 1122

Sequence: 1 MDQVSSPIVDINVTSEPC.....AACHGHLILGNFKNSGVSK 215

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 1122 | 100.0 | 215 | 18 AAW27408 | Inactive human CCR |
| 2 | 1122 | 100.0 | 215 | 20 AAW88238 | HIV-1 co-receptor |
| 3 | 958 | 85.4 | 184 | 18 AAW27406 | Inactive human CCR |
| 4 | 958 | 85.4 | 352 | 18 AAW27407 | Human CCR5, Homo |
| 5 | 958 | 85.4 | 352 | 18 AAW27123 | Human chemokine re |
| 6 | 958 | 85.4 | 352 | 19 AAW23835 | Human CC chemokine |
| 7 | 958 | 85.4 | 352 | 20 AAW88232 | HIV-1 co-receptor |
| 8 | 958 | 85.4 | 352 | 22 AAB56342 | Non-endogenous hum |
| 9 | 958 | 85.4 | 352 | 22 AAG80111 | Human CCR5 protein |
| 10 | 958 | 85.4 | 352 | 22 AAB82948 | Human HIV-1 co-rec |

| | | | | | |
|----|-------|------|-----|-------------|---------------------|
| 11 | 958 | 85.4 | 352 | 22 AAE07039 | Human G-protein ch |
| 12 | 958 | 85.4 | 352 | 22 AAE07048 | Human G-protein ch |
| 13 | 958 | 85.4 | 352 | 22 AAB83354 | Human CCR5 protein |
| 14 | 958 | 85.4 | 352 | 22 AAE04321 | Human chemokine re |
| 15 | 958 | 85.4 | 352 | 22 AAB46858 | Human HDGRI10 prot |
| 16 | 958 | 85.4 | 352 | 23 AAO97152 | Human G-protein ch |
| 17 | 958 | 85.4 | 352 | 23 ABB08343 | Human chemokine (C |
| 18 | 958 | 85.4 | 352 | 23 AAM52828 | Human CC chemokine |
| 19 | 958 | 85.4 | 439 | 20 AAY41280 | Fusion protein con |
| 20 | 953 | 84.9 | 371 | 19 AAW23834 | Human CC chemokine |
| 21 | 952 | 84.8 | 352 | 23 AAM52829 | Human CCR5 Gln 55 |
| 22 | 943 | 84.0 | 352 | 18 AAW07602 | Human G-protein ch |
| 23 | 943 | 84.0 | 352 | 21 AAY80128 | Human G-protein ch |
| 24 | 943 | 84.0 | 352 | 22 AAE07037 | Human G-protein ch |
| 25 | 943 | 84.0 | 352 | 22 AAE07046 | Human G-protein ch |
| 26 | 943 | 84.0 | 352 | 23 AAU97150 | Human G-protein ch |
| 27 | 936 | 83.4 | 352 | 22 AAG79089 | Amino acid sequenc |
| 28 | 933 | 83.2 | 352 | 18 AAW27125 | Macaque chemokine |
| 29 | 839 | 74.8 | 332 | 18 AAW26766 | Human chemokine re |
| 30 | 776 | 69.2 | 354 | 19 AAW54037 | Mouse CC-CKR5 prot |
| 31 | 695 | 61.9 | 360 | 22 AAU07613 | Human CCR2-641 pol |
| 32 | 694 | 61.9 | 360 | 16 AAR79166 | Human monocyte che |
| 33 | 694 | 61.9 | 360 | 18 AAW35833 | Non-endogenous hum |
| 34 | 694 | 61.9 | 360 | 22 ABB56340 | Human wild-type CC |
| 35 | 694 | 61.9 | 360 | 22 AAG07614 | Human monocyte che |
| 36 | 694 | 61.9 | 374 | 16 AAR79165 | Human CCR2a protei |
| 37 | 694 | 61.9 | 374 | 22 AAG80107 | Human MCP-1 recept |
| 38 | 694 | 61.9 | 329 | 22 AAB46859 | C-C chemokine rece |
| 39 | 608.5 | 54.2 | 329 | 22 AAB46859 | Human MIP-1 alpha/R |
| 40 | 598 | 53.3 | 355 | 15 AAR52749 | Human MIP-1 alpha/R |
| 41 | 598 | 53.3 | 355 | 18 AAW26588 | Human CC-chemokine |
| 42 | 598 | 53.3 | 355 | 18 AAW25751 | Rat CC chemokine r |
| 43 | 598 | 53.3 | 355 | 21 AAB20571 | Chemokine receptor |
| 44 | 565.5 | 50.4 | 355 | 18 AAW29179 | |
| 45 | 547.5 | 48.8 | 360 | 17 AAR99274 | |

ALIGNMENTS

RESULT 1

AAW27408
ID AAW27408 standard; Protein; 215 AA.

XX
AC AAW27408;

DT 14-APR-1998 (first entry)

DE Inactive human CCR5.

XX
KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.

OS Homo sapiens.

XX
PN WO9732019-A2.

XX
PD 04-SEP-1997.

XX
PF 28-FEB-1997; 97WO-BE00023.

XX
PR 06-AUG-1996; 96EP-0870102.

XX
PR 01-MAR-1996; 96EP-0870021.

XX
PA (EURO-) EUROSREEN SA.

XX
PI Libert F, Parmentier M, Samson M, Vassart G;

XX
DR WPI; 1997-479829/44.

```

DR N-PSDB; AAT90118.
XX Active and inactive forms of human CC chemokine receptor CCR-5 -
PT useful to diagnose, prevent and/or treat inflammatory disorders,
PT autoimmune disease and viral infection
XX Claim 7; Fig 1d-e; 94pp; English.
XX The present sequence is an inactive human CC (Cys-Cys)
CC chemokine receptor 5 (CCR5), which lacks the last 3 transmembrane
CC regions and the regions involved in G protein-coupling. CCR5 or
CC its cDNA can be used to diagnose, treat and/or prevent inflammatory
CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,
CC idiopathic pulmonary fibrosis and psoriasis, viral infections,
CC especially human immunodeficiency virus type 1 or type 2 (HIV-1 or
CC HIV-2) infection, cancer, atherosclerosis and autoimmune disorders.
CC Subjects that express the inactive receptor have a predisposition,
CC or resistance to HIV-1 and/or HIV-2.
XX Sequence 215 AA;

Query Match 100.0%; Score 1122; DB 18; Length 215;
Best Local Similarity 100.0%; Pred. No. 1.4e-125;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCCOLLTGLYFIFGFSGIFFT 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCCOLLTGLYFIFGFSGIFFT 120
QY 121 LLTIDRYLAVHVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DB 121 LLTIDRYLAVHVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HFPYIKDHLGAGPAAACHGHLILLGNPKNSASVSK 215
DB 181 HFPYIKDHLGAGPAAACHGHLILLGNPKNSASVSK 215

RESULT 2
AAW88238
ID AAW88238 standard; Protein; 215 AA.
AC AAW88238;
XX 15-MAR-1999 (first entry)
XX HIV-1 co-receptor CCR5 variant CCR5-delta32.
DE HIV-1 co-receptor CCR5 variant CCR5-delta32.
XX HIV-1; CCR5; CCR5-delta32; co-receptor; infection; diagnosis; AIDS;
KW gene therapy; human.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Domain 32..56
FT Domain /note= "transmembrane domain 1"
FT Domain 67..87
FT Domain /note= "transmembrane domain 2"
FT Domain 103..124
FT Domain /note= "transmembrane domain 3"
FT Domain 142..167
FT Domain /note= "transmembrane domain 4"
XX W09854317-A1.
PN 03-DEC-1998.
PD 29-MAY-1998; 98WO-EP03437.
XX

```

```

PR 30-MAY-1997; 97US-0048057.
XX (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
XX Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
XX WPI; 1999-059835/05.
XX N-PSDB; AAW84159.
XX New CCR5 variant protein of the HIV-1 co-receptor - useful in
PT developing resistance of CCR5-expressing cells to HIV-1 infection
XX Disclosure; Page 38-39; 55pp; English.
CC This is the amino acid sequence of a CCR5 variant protein,
CC designated CCR5-delta32, that includes the first 4 transmembrane
CC domains of wild-type CCR5 (see AAW88232), but lacks transmembrane
CC domains 5-7. CCR5 serves as a co-receptor for infection by
CC macrophage-tropic (M-tropic) strains of HIV-1. Individuals
CC homozygous for the CCR5-delta32 mutation are resistant to HIV-1
CC infection, but heterozygous individuals are susceptible. The
CC invention additionally relates to the identification of variant
CC CCR5s303 (see AAW88231), which lacks transmembrane domains 3-7 of
CC CCR5. The detection of CCR5 variants may be used to identify
CC individuals at lower risk of infection relative to the general
CC population who, if infected, may exhibit slower progression to
CC AIDS. Probes and primers (see AAW84127-36) are provided for use in
CC diagnostic methods for detecting the presence of such variants. A
CC method is provided for inhibiting HIV-1 infection of a cell
CC expressing the CCR5 receptor. This involves introducing a nucleic
CC acid encoding a CCR5 variant into the cell, thereby reducing the
CC number of functional CCR5 molecules present on the cell surface.
XX Sequence 215 AA;

Query Match 100.0%; Score 1122; DB 20; Length 215;
Best Local Similarity 100.0%; Pred. No. 1.4e-125;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMVLILLINCKR 60
QY 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCCOLLTGLYFIFGFSGIFFT 120
DB 61 LKSMTDIYLLNLAIISDLFFLLTPVFWAHYAAQWDFGNTMCCOLLTGLYFIFGFSGIFFT 120
QY 121 LLTIDRYLAVHVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
DB 121 LLTIDRYLAVHVFALKARTVTFGVVTSVITWVAVFASLPGLIFTRSQEGLHYTCSS 180
QY 181 HFPYIKDHLGAGPAAACHGHLILLGNPKNSASVSK 215
DB 181 HFPYIKDHLGAGPAAACHGHLILLGNPKNSASVSK 215

RESULT 3
AAW27406
ID AAW27406 standard; Protein; 184 AA.
XX AC AAW27406;
XX 14-APR-1998 (first entry)
XX Inactive human CCR5.
XX Inactive; human Cys-Cys chemokine receptor-5; CCR5;
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
KW predisposition; resistance; diagnosis; treatment; prevention;
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
KW cancer; atherosclerosis; autoimmune disorder.
XX

```

OS Homo sapiens.
 XX W09732019-A2.
 XX PD 04-SEP-1997.
 XX PF 28-FEB-1997; 97WO-BE00023.
 XX PR 06-AUG-1996; 96EP-0870102.
 XX PR 01-MAR-1996; 96EP-0870021.
 XX PA (EURO-) EUROSREEN SA.
 XX PI Libert F, Parmentier M, Samson M, Vassart G;
 XX WPI; 1997-479829/44.
 XX DR N-PSDB; AAT90116.
 XX PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 XX useful to diagnose, prevent and/or treat inflammatory disorders,
 XX autoimmune disease and viral infection
 XX PS Claim 1; Fig 1a; 94pp; English.
 XX CC The present sequence is an inactive human CC (Cys-Cys)
 CC chemokine receptor 5 (CCR5), which is not a receptor of human
 CC immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2). CCR5 or
 CC its cDNA can be used to diagnose, treat and/or prevent inflammatory
 CC diseases, e.g. rheumatoid arthritis, glomerulonephritis, asthma,
 CC idiopathic pulmonary fibrosis and psoriasis, viral infections,
 CC especially HIV-1 or HIV-2 infection, cancer, atherosclerosis and
 CC autoimmune disorders. Subjects that express the inactive receptor
 CC have a predisposition, or resistance to HIV-1 and/or HIV-2.
 XX SQ Sequence 184 AA;
 Query Match 85.4%; Score 958; DB 18; Length 184;
 Best Local Similarity 100.0%; Pred. No. 4.4e-106;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDYQSSPIYDINVTSEPCOKINVKQIAARLLPLYSLVFIFGFGNMLVILLINCKR 60
 Db 1 MDYQSSPIYDINVTSEPCOKINVKQIAARLLPLYSLVFIFGFGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCOLLTGLYFIFGSGIFII 120
 Db 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCOLLTGLYFIFGSGIFII 120
 QY 121 LITDRIYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 Db 121 LITDRIYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 Db 181 HPFY 184
 RESULT 4
 AAW27407
 ID AAW27407 standard; Protein; 352 AA.
 XX AC AAW27407;
 XX DT 14-APR-1998 (first entry)
 XX DE Human CCR5.
 XX KW Human Cys-Cys chemokine receptor 5; CCR5;
 KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;
 KW diagnosis; treatment; prevention;
 KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;
 KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;
 KW cancer; atherosclerosis; autoimmune disorder.

XX OS Homo sapiens.
 XX PN W09732019-A2.
 XX XX PD 04-SEP-1997.
 XX PF 28-FEB-1997; 97WO-BE00023.
 XX PR 06-AUG-1996; 96EP-0870102.
 XX PR 01-MAR-1996; 96EP-0870021.
 XX PA (EURO-) EUROSREEN SA.
 XX PI Libert F, Parmentier M, Samson M, Vassart G;
 XX WPI; 1997-479829/44.
 XX DR N-PSDB; AAT90117.
 XX PT Active and inactive forms of human CC chemokine receptor CCR-5 -
 XX useful to diagnose, prevent and/or treat inflammatory disorders,
 XX autoimmune disease and viral infection
 XX PS Claim 4; Fig 1b-c; 94pp; English.
 XX CC The present sequence is human CC (Cys-Cys) chemokine receptor
 CC 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES
 CC chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),
 CC MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product
 CC alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of
 CC human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).
 CC CCR5 or its cDNA can be used to diagnose, treat and/or prevent
 CC inflammatory diseases, e.g. rheumatoid arthritis,
 CC glomerulonephritis, asthma, idiopathic pulmonary fibrosis and
 CC psoriasis, viral infections, especially HIV-1 or HIV-2 infection,
 CC cancer, atherosclerosis and autoimmune disorders.
 XX SQ Sequence 352 AA;
 Query Match 85.4%; Score 958; DB 18; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDYQSSPIYDINVTSEPCOKINVKQIAARLLPLYSLVFIFGFGNMLVILLINCKR 60
 Db 1 MDYQSSPIYDINVTSEPCOKINVKQIAARLLPLYSLVFIFGFGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCOLLTGLYFIFGSGIFII 120
 Db 61 LKSMTDIYLLNLAISDLFFLLTPFWAHYAAQWDFGNTMCOLLTGLYFIFGSGIFII 120
 QY 121 LITDRIYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 Db 121 LITDRIYLVAVHAFKARTVFGVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HPFY 184
 Db 181 HPFY 184
 RESULT 5
 AAW27123
 ID AAW27123 standard; Protein; 352 AA.
 XX AC AAW27123;
 XX DT 14-DEC-1997 (first entry)
 XX DE Human chemokine receptor 88C.
 XX KW Chemokine receptor 88C; atherosclerosis; rheumatoid arthritis;
 KW tumour; asthma; viral infection; AIDS; inflammation;
 KW autoimmune disease; therapy; diagnosis; leukocyte trafficking;

G protein coupled receptor; ligand; modulator; antibody; human.

Homo sapiens.

Key Location/Qualifiers

Domain 1..32

Domain /label= Extracellular_domain

Domain 56..67

Domain /label= Intracellular_domain

Domain 89..112

Domain /label= Extracellular_domain

Domain 125..145

Domain /label= Intracellular_domain

Domain 166..191

Domain /label= Extracellular_domain

Domain 213..235

Domain /label= Intracellular_domain

Domain 259..280

Domain /label= Extracellular_domain

Domain 301..352

Domain /label= Intracellular_domain

WO9722698-A2.

26-JUN-1997.

20-DEC-1996; 96WO-US20759.

07-JUN-1996; 96US-0661393.

20-DEC-1995; 95US-0575967.

(ICOS-) ICOS CORP.

Gray PW, Raport CJ, Schweickart VL;

WPI: 1997-341689/31.

N-PSDB; AAT85161.

New nucleic acid encoding chemokine receptors 88-2B and 88C - used to modulate leukocyte trafficking, e.g. for treatment of inflammation, tumours, viral infections, autoimmune diseases, etc.

Claim 16; Page 47-48; 65pp; English.

This polypeptide sequence comprises novel human chemokine receptor 88C, a G protein coupled receptor that is involved in leukocyte trafficking. Its amino sequence was deduced from a cDNA clone (AAT85161) isolated from a macrophage library. It shows 62% identity to CCR1. Chemokine receptor 88-2B (see AAW27124) has also been identified. 88C and 88-2B receptors and their polypeptide fragments can be produced in transformed host cells. The receptors, peptides comprising one or more of the extracellular or intracellular domains, and anti-receptor antibodies can be used to modulate receptor activities, particularly ligand and G protein binding, and are potentially potentially useful in the treatment of atherosclerosis, rheumatoid arthritis, tumours, asthma, viral infection, AIDS, inflammatory conditions, pathological immune response, abnormal haematopoietic processes etc.

Sequence 352 AA;

Query Match 85.4%; Score 958; DB 18; Length 352;

Best Local Similarity 100.0%; Pred. No. 1e-105;

Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSPIDYNYTSEPCQKINVKQIAARLLPPLYSILVFYGVGNMLVILINCKR 60

DB 1 MDYQVSPIDYNYTSEPCQKINVKQIAARLLPPLYSILVFYGVGNMLVILINCKR 60

QY 61 LKSMTDYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMCQLLTGLTYFGFFSIFII 120

DB 61 LKSMTDYLLNLAISDLFLLTVPFWAHYAAQWDFGNTMCQLLTGLTYFGFFSIFII 120

QY 121 LLTIDRYLAVHAFVFAKARVTFGVTVSWVAVFASLPGLIFTRSQEGLHYTCSS 180

DB 121 LLTIDRYLAVHAFVFAKARVTFGVTVSWVAVFASLPGLIFTRSQEGLHYTCSS 180

QY 181 HFPY 184

DB 181 HFPY 184

RESULT 6

AAW23835

ID AAW23835 standard; Protein; 352 AA.

XX AC AAW23835;

AC AAW23835;

DT 08-JUN-1998 (first entry)

XX DE Human CC chemokine receptor 5 (CCR5).

XX KW CC chemokine receptor 5; CCR5; G-protein coupled receptor;

KW human immunodeficiency virus; HIV; CD4; AIDS; therapy;

KW transgenic animal.

XX OS Homo sapiens.

XX FH Key

FT Domain Location/Qualifiers

FT 29..55

FT /label= I

FT /note= "transmembrane domain"

FT 109..120

FT Region

FT /note= "extracellular loop-1 (Claim 19)"

FT 104..126

FT Domain

FT /label= III

FT /note= "transmembrane domain"

FT 143..171

FT Domain

FT /label= IV

FT /note= "transmembrane domain"

FT 187..210

FT Region

FT /note= "extracellular loop-2 (Claim 19)"

FT 194..219

FT Domain

FT /label= V

FT /note= "transmembrane domain"

FT 238..258

FT Domain

FT /label= VI

FT /note= "transmembrane domain"

FT 261..276

FT Region

FT /note= "extracellular loop-3 (Claim 19)"

FT 277..300

FT Domain

FT /label= VII

FT /note= "transmembrane domain"

WO9745543-A2.

04-DEC-1997.

28-MAY-1997; 97WO-US09586.

28-MAY-1996; 96US-0018508.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Alkhatib G, Berger EA, Broder CC, Combadiere C;

PI Feng Y, Kennedy PE, Murphy PM;

XX WPI: 1998-032650/03.

DR N-PSDB; AAT76920.

XX CC Chemokine receptor 5 polypeptide - used to inhibit membrane

XX fusion between HIV and a target cell

XX Claim 68; Fig 1C; 70pp; English.

XX This protein sequence comprises of a novel human macrophage-selective

CC chemokine receptor that has been designated CCR5. The sequence was deduced from an isolated cDNA clone (see AAR76920). An Alai27uen variant (see W238340) of CCR5 was also identified. The susceptibility of human macrophages to HIV infection depends on cell surface expression of CD4 and CCR5. CCR5 is a member of the 7-transmembrane superfamily of G-protein coupled cell surface molecules. It plays an essential role in the membrane fusion step of infection by some HIV isolates. The establishment of stable, non-human cell lines and transgenic mammals having cells that coexpress human CD4 and CCR5 provides valuable tools for research of HIV infection. Antibodies that bind to CCR5, CCR5 variants, and CCR5-binding agents capable of blocking membrane fusion between HIV and target cells represent potential anti-HIV therapeutics for macrophage tropic strains of HIV.

XX
SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 19; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-105;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMVDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
DB 61 LKSMVDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
QY 121 LITIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQKESGLHYTCSS 180
DB 121 LITIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQKESGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 7
AAW88232
ID AAW88232 standard; Protein; 352 AA.

XX
AC AAW88232;
DT 15-MAR-1999 (first entry)
DE HIV-1 co-receptor CCR5.
XX
KW HIV-1; CCR5; CCR5m303; co-receptor; infection; diagnosis; AIDS;
KW gene therapy; human.
XX
OS Homo sapiens.

Key Location/Qualifiers
FH Domain 32..56
FT /note= "transmembrane domain 1"
FT Domain 67..87
FT /note= "transmembrane domain 2"
FT Misc-difference 101
FT /note= "corresponds to TGR (Cys) in wild-type CCR5,
FT TGA (Stop) in CCR5m303"
FT Domain 103..124
FT /note= "transmembrane domain 3"
FT Domain 142..167
FT /note= "transmembrane domain 4"
FT Domain 200..223
FT /note= "transmembrane domain 5"
FT Domain 236..260
FT /note= "transmembrane domain 6"
FT Domain 275..301
FT /note= "transmembrane domain 7"

XX
PN W09854317-A1.

XX
PD 03-DEC-1998.
XX
PF 29-MAY-1998; 98MO-EP03437.
XX
PR 30-MAY-1997; 97US-0048057.
XX
PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
XX
PI Arenzana Siesdedos F, Beretta A, Braun J, Quillent C;
XX
DR WPI: 1999-059835/05.
XX
DR N-PSDB; AAV841126.
XX
PT New CCR5 variant protein of the HIV-1 co-receptor - useful in
XX developing resistance of CCR5-expressing cells to HIV-1 infection
XX
PS Disclosure; Page 34-35; 55pp; English.
XX
CC This is the amino acid sequence of wild-type human CCR5, which
CC serves as a co-receptor for infection by macrophage-tropic
CC (M-tropic) strains of HIV-1. The invention relates to the
CC identification of a CCR5 variant (see AAW88231), designated CCR5m303,
CC comprising the first two transmembrane domains of wild-type CCR5,
CC but lacking transmembrane domains 3-7. The presence of the
CC CCR5m303 variant with the wild type CCR5 allele shows a positive
CC correlation with resistance to infection with M-tropic HIV-1
CC strains, and may indicate slower progression of the disease. The
CC detection of CCR5 variants may be used to identify individuals at
CC lower risk of infection relative to the general population who, if
CC infected, may exhibit slower progression to AIDS. Probes and
CC primers (see AAW84127-36) are provided for use in diagnostic methods
CC for detecting the presence of such variants. A method is provided
CC for inhibiting HIV-1 infection of a cell expressing the CCR5
CC receptor. This involves introducing a nucleic acid encoding a CCR5
CC variant into the cell, thereby reducing the number of functional
CC CCR5 molecules present on the cell surface.

XX
SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 20; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-105;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
DB 1 MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVIFGFGVGNMVLILINCKR 60
QY 61 LKSMVDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
DB 61 LKSMVDIYLLNLAIISDLFLLTPFWAHYAAQWDFGNTMCGLLTGLYFGFGSIFPII 120
QY 121 LITIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQKESGLHYTCSS 180
DB 121 LITIDRYLAVHVAFAKARTVFGVTSVITWVAVFASLPGLIIFTRSQKESGLHYTCSS 180
QY 181 HPFY 184
DB 181 HPFY 184

RESULT 8
ABB56342
ID ABB56342 standard; Protein; 352 AA.
XX
AC ABB56342;
XX
DT 18-FEB-2002 (first entry)
XX
DE Non-endogenous human GPCR protein, SEQ ID NO: 477.
KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KW constitutively activated GPCR; agonist; disease.

| | |
|---|----|
| CCRS5; chemokine; co-receptor; human immunodeficiency virus type 1; HIV-1; Infection; therapy; vaccine; anti-HIV-1. | XX |
| Homo sapiens. | XX |
| Key Location/Qualifiers | XX |
| Binding-site 2.18 | XX |
| /note= "binds to HIV-1 gp120" | XX |
| WO200164710-A2. | XX |
| 07-SEP-2001. | XX |
| 28-FEB-2001; 2001WO-US06699. | XX |
| 29-FEB-2000; 2000US-185667P. | XX |
| 19-MAY-2000; 2000US-205839P. | XX |
| 07-FEB-2001; 2001US-267231P. | XX |
| (PROG-) PROGENICS PHARM INC. | XX |
| (AARO-) AARON DIAMOND AIDS RES CENT. | XX |
| Dragic T, Olson WC; | XX |
| WPI; 2001-611273/70. | XX |
| N-PSDB; AAH26903. | XX |
| Novel compounds comprising specific amino acids within CCR5 (HIV 1 co-receptor) amino terminal domain including negatively charged and two sulfated tyrosine residues is useful for treating HIV infection in humans | XX |
| Claim 1; Page 30; 163pp; English. | XX |
| The present sequence is that of human HIV-1 co-receptor CCR5. Amino acids 2-18 in the N-terminal region of CCR5 comprise an HIV-1 gp120-binding site that determines the specificity of the interaction between CCR5 and HIV-1 gp120. Post-translational sulfation of the tyrosine residues in the CCR5 N-terminus is required for gp120 binding and may critically modulate the susceptibility of target cells to HIV-1 infection in vivo. The invention provides claimed sulfated peptides (see AAB82947) that are based on the CCR5 N-terminal region and which are effective for inhibiting HIV-1 binding to CCR5. These peptides are used in claimed methods of inhibiting HIV infection of CD4+ cells, of preventing CD4+ cells from becoming infected with HIV, of treating a subject whose CD4+ cells are infected with HIV, and of identifying an agent which inhibits binding of a CCR5 ligand to a CCR5 receptor. The methods may be carried out in a subject, especially a human, infected (therapeutic method), not infected with HIV (prophylactic method), or in a subject who is not infected with, but has been exposed to, HIV. | XX |
| Sequence 352 AA; | XX |
| Query Match 85.4%; Score 958; DB 22; Length 352; | |
| Best Local Similarity 100.0%; Pred. No. 1e-105; | |
| Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps | |
| QY 1 MDYQSSPYDIINYYTSPCKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILLINCKR 61 | |
| Db 1 MDYQSSPYDIINYYTSPCKINVKQIAARLLPPLYSLVFIFGFGVGNMLVILLINCKR 61 | |
| QY 61 LKSMYDYLNLALSDLFLLTPVFAHVAQAQNDFGNTWCOLTGLYFTGFPFGIFPII 1 | |
| Db 61 LKSMYDYLNLALSDLFLLTPVFAHVAQAQNDFGNTWCOLTGLYFTGFPFGIFPII 1 | |
| QY 121 LKSIDRYLVAVHAFKARVTFGVVTSVITWWAVFASLPGLIFTRSOKEGLHWTCSS 1 | |
| Db 121 LKSIDRYLVAVHAFKARVTFGVVTSVITWWAVFASLPGLIFTRSOKEGLHWTCSS 1 | |
| QY 181 HPFY 184 | |
| XX | |

SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPFMAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPFPFMAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 QY 121 LLTIDRYLAVVHAFKARTVTVGVVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 DB 121 LLTIDRYLAVVHAFKARTVTVGVVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 12

AAE07048
 ID AAE07048 standard; Protein; 352 AA.

XX AC AAE07048;

XX DT 16-OCT-2001 (first entry)

XX DE Human G-protein chemokine receptor (CCR5) HDGNR10 protein #2.
 KW Human; G-protein chemokine receptor; CCR5; HDGNR10; inflammation; HIV;
 KW human immunodeficiency virus; antimicrobial; vasodilator; vulnery;
 KW cystostatic; immunosuppressive; neutrotic; neuroprotective; gene therapy;
 KW neurodegenerative disorder; Kaposi's sarcoma; autoimmune disease;
 KW rheumatoid arthritis; cancer; breast; ovary; adrenal gland; bone marrow;
 KW gastrointestinal tract; lung; liver; immune disorder; Addison's disease;
 KW haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; allergy;
 KW multiple sclerosis; ulcerative colitis; Crohn's disease; wound healing;
 KW cardiovascular disorder; myocardial ischaemia.

XX OS Homo sapiens.

XX PN WO200158916-A2.

XX PD 16-AUG-2001.

XX PF 09-FEB-2001; 2001WO-US04153.

XX PR 09-FEB-2000; 2000US-0181258.

XX PR 09-MAR-2000; 2000US-0187999.

XX PR 22-SEP-2000; 2000US-0234336.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Roschke V, Li Y, Ruben SM;

XX DR WPI: 2001-488966/53.

XX DR N-PSDB; AAD13299.

XX PT Isolated nucleic acid encoding a human G-protein chemokine receptor
 PT (CCR5) HDGNR10 polypeptide, useful for preventing or treating
 PT autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 PT disorders and neurodegenerative disorders -
 XX Example 40; Page 504-505; 518pp; English.

XX CC The invention relates to human G-protein chemokine receptor (CCR5)
 CC HDGNR10 polypeptides and polynucleotides. CCR5 HDGNR10 antibodies
 CC are useful for treating, preventing or ameliorating a disease or disorder
 CC associated with inflammation, defective or aberrant chemotaxis of immune

CC cells, HIV infection (such as Pneumocystis carinii pneumonia or Kaposi's
 CC sarcoma) or defective or aberrant T-cell antigen presenting cell
 CC interaction. The disease or disorder may also be an infectious disease
 CC (e.g. a viral infection such as an early stage HIV infection, a
 CC cytomegalovirus infection, or a poxvirus infection), an autoimmune
 CC disease (e.g. rheumatoid arthritis) or a neurodegenerative disorder. The
 CC disease or disorder may be associated with aberrant CCR5 expression, lack
 CC of CCR5 function, aberrant CCR5 ligand expression, or lack of CCR5 ligand
 CC function. CCR5 HDGNR10 protein is used as a food additive or preservative
 CC to increase or decrease storage capabilities. CCR5 HDGNR10 DNA are useful
 CC for chromosome identification and in gene therapy. CCR5 HDGNR10 DNA, RNA,
 CC protein, antibodies, agonists and antagonists are also useful in the
 CC diagnosis, treatment and prevention of cancer (breast, ovary, adrenal
 CC gland, bone, bone marrow, gastrointestinal tract, liver, lung,
 CC urogenital); immune disorders (Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis); cardiovascular disorders (myocardial ischaemia) and
 CC wound healing. The present sequence is human CCR5 HDGNR10 protein.

XX SQ Sequence 352 AA;

Query Match 85.4%; Score 958; DB 22; Length 352;

Best Local Similarity 100.0%; Pred. No. 1e-105;

Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
 DB 1 MDYQVSSPIYDINTYTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNMLVILLINCKR 60
 QY 61 LKSMTDIYLLNLAIISDLFFLLTPFPFMAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLAIISDLFFLLTPFPFMAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 QY 121 LLTIDRYLAVVHAFKARTVTVGVVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 DB 121 LLTIDRYLAVVHAFKARTVTVGVVTSVITWVAVFASLPGLIIFRSQKGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184

RESULT 13

AAE83354
 ID AAE83354 standard; Protein; 352 AA.

XX AC AAE83354;

XX DT 09-OCT-2001 (first entry)

XX DE Human CCR5 protein sequence.

XX KW Chemotactic chemokine receptor 5; gp120; CD4; therapy; HIV; CCR5;
 KW human immunodeficiency virus; anti-inflammatory disease; human.
 XX OS Homo sapiens.

XX PN EPI118858-A2.

XX PD 25-JUL-2001.

XX PF 03-JAN-2001; 2001EP-0300020.

XX PR 12-JAN-2000; 2000GB-0000659.

XX PR 12-JAN-2000; 2000GB-0000661.

XX PR 12-JAN-2000; 2000GB-0000663.

XX PA (PFIZ) PFIZER LTD.

XX PA (PFIZ) PFIZER INC.

XX PI Dobbs S, Perros M, Rickett GA;

XX XX

DR WPI: 2001-477085/52.
 DR N-PSDB; AAF87099.
 XX
 PT Determining if an agent can modulate CCR5-gp120 interaction, comprises
 PT incubating the agent with CCR5 and gp120 and determining if the agent
 PT modulates the interaction
 XX
 PS Claim 1: Page 110; 113pp; English.
 XX
 CC This sequence represents the human CCR5 protein sequence.
 CC The invention relates to a method for determining whether an agent is
 CC capable of modulating the interaction of chemotactic chemokine receptor 5
 CC (CCR5) with gp120, comprising incubating the agent with CCR5 and gp120
 CC and determining whether the agent modulates the interaction, where gp120
 CC is associated with CD4, and where the interaction is a low affinity
 CC binding. The method is used to identify an agent capable of modulating
 CC the interaction of CCR5 with gp120. An agent identified by the method
 CC is used to prepare a pharmaceutical composition for the treatment of a
 CC disease or condition associated with CCR5 and gp120 interaction, to treat
 CC a subject with a disease or condition associated with CCR5 and gp120
 CC interaction, and for preparing a pharmaceutical for treating human
 CC immunodeficiency virus (HIV). It can also be used to treat anti-inflammatory
 CC diseases. The method is commercially useful, amenable to high throughput
 CC screening, and detects interaction of gp120 with cells expressing only
 CC CCR5.
 XX
 SQ Sequence 352 AA;
 Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPLYSILVFIFGVGNMVLILLINCKR 60
 DB 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPLYSILVFIFGVGNMVLILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLAISDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 QY 121 LTTIDRYLAVVHAFVFAKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTIDRYLAVVHAFVFAKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184
 RESULT 14
 AAE04321
 ID AAE04321 standard; Protein; 352 AA.
 XX
 AC AAE04321;
 XX
 XX 04-SEP-2001 (first entry)
 XX
 DE Human chemokine receptor (CCR), CC-CR-5 related protein #2.
 XX
 KW Human; transformed mammalian cell; CD4; reporter gene; translocation;
 KW human immuno deficiency virus; HIV; long terminal repeat; LTR; therapy;
 KW chemokine receptor; CCR; cellular dysfunction; HIV infection; cofactor;
 KW CC-CR-5; envelope glycoprotein; anti-HIV.
 XX
 OS Homo sapiens.
 XX
 XX US6258527-B1.
 XX
 PD 10-JUL-2001.
 XX
 XX 21-MAY-1997; 97US-0861105.
 XX
 XX 20-MAY-1996; 96US-0017157.

PR 19-JUN-1996; 96US-0020043.
 PR 19-MAY-1997; 97US-0858660.
 XX
 PA (AARO-) AARON DIAMOND AIDS RES CENT.
 PA (UINY) UNIV NEW YORK STATE.
 XX
 PI Littman DR, Deng H, Ellmeier W, Landau NR, Liu R;
 XX
 DR WPI: 2001-417127/44.
 DR N-PSDB; AAD08577.
 XX
 PT Transformed mammalian cell (I) that contains a CD4 gene, reporter gene
 PT and HIV LTR for identification of drugs and antibodies for treatment of
 PT HIV -
 XX
 PS Disclosure; Column 47-50; 37pp; English.
 XX
 CC The present invention relates to a transformed mammalian cell that
 CC contains a gene encoding CD4, a construct encoding a reporter gene under
 CC the regulation of an human immuno deficiency virus (HIV) long terminal
 CC repeat (LTR) and that has been transduced with a vector encoding a
 CC human chemokine receptor (CCR) where the CD4 and the CCR are present on
 CC the cell surface of transformed mammalian cell. The invention is useful
 CC for identifying drugs or antibodies that interfere with the
 CC translocation of HIV into transformed mammalian cell or for identifying
 CC a human chemokine receptor that facilitates the infection of a
 CC particular HIV strain into the transformed mammalian cell. Compounds
 CC identified can be used to treat cellular dysfunction and to prevent or
 CC combat HIV infection. The present sequence is a human chemokine receptor
 CC (CCR), CC-CR-5 related protein. CC-CR-5 is the principal cofactor for
 CC entry mediated by the envelope glycoproteins of primary macrophage-tropic
 CC strains of HIV-1.
 XX
 SQ Sequence 352 AA;
 Query Match 85.4%; Score 958; DB 22; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1e-105;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPLYSILVFIFGVGNMVLILLINCKR 60
 DB 1 MDYQVSSPIYDINVTSEPCQKINVKQIAARLLPLYSILVFIFGVGNMVLILLINCKR 60
 QY 61 LKSMTDIYLLNLAISDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 DB 61 LKSMTDIYLLNLAISDLFFLLTPVFWAHYAAQWDFGNTMCQLLTGLYFIFGFSGIFPII 120
 QY 121 LTTIDRYLAVVHAFVFAKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 DB 121 LTTIDRYLAVVHAFVFAKARTVTFGVVTSVITWVAVFASLPGLIIFTRSQEGLHYTCSS 180
 QY 181 HFPY 184
 DB 181 HFPY 184
 RESULT 15
 AAB46858
 ID AAB46858 standard; Protein; 352 AA.
 XX
 AC AAB46858;
 XX
 DT 16-AUG-2001 (updated)
 DT 02-AUG-2001 (updated)
 DT 04-MAY-2001 (first entry)
 XX
 DE Human HDGMR10 protein.
 XX
 KW HDGMR10; human; G-protein chemokine receptor; antiinflammatory;
 KW immunomodulatory; anticoagulant; antiallergic; immunosuppressive;
 KW cytostatic; antiparasitic; antiposrotic; antirheumatic; antiarthritic;
 KW vasotropic; gene therapy; haematopoiesis; wound healing; coagulation;
 KW angiogenesis; solid tumour; infection; leukemia; growth factor activity;

Search completed: June 3, 2003, 19:22:34
Job time : 56.061 secs

Qy 181 HFPY 184

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 3, 2003, 18:49:55 ; Search time 4.22266 Seconds
(without alignments)
4506.655 Million cell updates/sec

Title: US-09-938-703-3_COPY_792_885

Perfect score: 164

Sequence: 1 ATTAAAGATGATCATCTGG.....CTCTGCTGCTGGTGTGGAAT 94

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 767038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=slp
-O=/cgn2_1/USPTO_pool/US0938703/runat_03062003_161410_23015/app_query.fasta_1.1870
-DB=Published_Applications_AA -QMT=frastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=ptt -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0938703@cgn.1.1.24 @runat_03062003_161410_23015
-NCPD=6 -ICPD=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|------------------------------------|
| 1 | 164 | 100.0 | 34 | 10 | US-09-938-719-13 Sequence 13, Appl |
| 2 | 164 | 100.0 | 34 | 10 | US-09-939-226-11 Sequence 13, Appl |
| 3 | 164 | 100.0 | 34 | 10 | US-09-938-703-13 Sequence 13, Appl |
| 4 | 164 | 100.0 | 215 | 10 | US-09-938-719-6 Sequence 6, Appl |

| | | | | | | |
|----|-----|-------|-----|----|-------------------|-------------------|
| 5 | 164 | 100.0 | 215 | 10 | US-09-938-226-6 | Sequence 6, Appl |
| 6 | 164 | 100.0 | 215 | 10 | US-09-938-703-6 | Sequence 6, Appl |
| 7 | 152 | 92.7 | 54 | 10 | US-09-938-719-11 | Sequence 11, Appl |
| 8 | 152 | 92.7 | 54 | 10 | US-09-939-226-11 | Sequence 11, Appl |
| 9 | 152 | 92.7 | 54 | 10 | US-09-938-703-11 | Sequence 11, Appl |
| 10 | 152 | 92.7 | 352 | 9 | US-10-232-686-2 | Sequence 2, Appl |
| 11 | 152 | 92.7 | 352 | 9 | US-10-086-814-1 | Sequence 1, Appl |
| 12 | 152 | 92.7 | 352 | 9 | US-09-734-221A-14 | Sequence 14, Appl |
| 13 | 152 | 92.7 | 352 | 10 | US-09-725-285-2 | Sequence 2, Appl |
| 14 | 152 | 92.7 | 352 | 10 | US-09-759-841-2 | Sequence 2, Appl |
| 15 | 152 | 92.7 | 352 | 10 | US-09-779-879A-2 | Sequence 2, Appl |
| 16 | 152 | 92.7 | 352 | 10 | US-09-779-879A-22 | Sequence 22, Appl |
| 17 | 152 | 92.7 | 352 | 10 | US-09-779-880A-2 | Sequence 2, Appl |
| 18 | 152 | 92.7 | 352 | 10 | US-09-779-880A-22 | Sequence 22, Appl |
| 19 | 152 | 92.7 | 352 | 10 | US-09-813-653-15 | Sequence 15, Appl |
| 20 | 152 | 92.7 | 352 | 10 | US-09-813-653-17 | Sequence 17, Appl |
| 21 | 152 | 92.7 | 352 | 10 | US-09-796-202-1 | Sequence 1, Appl |
| 22 | 152 | 92.7 | 352 | 10 | US-09-195-662A-2 | Sequence 2, Appl |
| 23 | 152 | 92.7 | 352 | 10 | US-09-339-912A-2 | Sequence 2, Appl |
| 24 | 152 | 92.7 | 352 | 10 | US-09-938-719-5 | Sequence 5, Appl |
| 25 | 152 | 92.7 | 352 | 10 | US-09-939-226-5 | Sequence 5, Appl |
| 26 | 152 | 92.7 | 352 | 10 | US-09-938-703-5 | Sequence 5, Appl |
| 27 | 152 | 92.7 | 352 | 10 | US-09-502-783A-2 | Sequence 2, Appl |
| 28 | 152 | 92.7 | 352 | 12 | US-10-106-623-2 | Sequence 2, Appl |
| 29 | 149 | 90.9 | 352 | 12 | US-10-106-623-20 | Sequence 20, Appl |
| 30 | 134 | 81.7 | 329 | 10 | US-09-725-285-9 | Sequence 9, Appl |
| 31 | 134 | 81.7 | 329 | 10 | US-09-195-662A-9 | Sequence 9, Appl |
| 32 | 134 | 81.7 | 329 | 10 | US-09-339-912A-9 | Sequence 9, Appl |
| 33 | 134 | 81.7 | 329 | 10 | US-09-502-783A-9 | Sequence 9, Appl |
| 34 | 134 | 81.7 | 344 | 9 | US-10-232-686-9 | Sequence 9, Appl |
| 35 | 134 | 81.7 | 344 | 10 | US-09-779-879A-9 | Sequence 9, Appl |
| 36 | 134 | 81.7 | 344 | 10 | US-09-779-880A-9 | Sequence 9, Appl |
| 37 | 134 | 81.7 | 347 | 10 | US-09-104-792-3 | Sequence 3, Appl |
| 38 | 134 | 81.7 | 360 | 10 | US-09-131-827A-2 | Sequence 2, Appl |
| 39 | 134 | 81.7 | 360 | 10 | US-09-131-827A-20 | Sequence 20, Appl |
| 40 | 134 | 81.7 | 360 | 10 | US-09-938-719-7 | Sequence 7, Appl |
| 41 | 134 | 81.7 | 360 | 10 | US-09-939-226-7 | Sequence 7, Appl |
| 42 | 134 | 81.7 | 360 | 10 | US-09-938-703-7 | Sequence 7, Appl |
| 43 | 109 | 66.5 | 332 | 9 | US-10-001-835-140 | Sequence 140, App |
| 44 | 109 | 66.5 | 355 | 9 | US-09-922-895-1 | Sequence 1, Appl |
| 45 | 109 | 66.5 | 355 | 10 | US-09-931-381A-16 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1
US-09-938-719-13
; Sequence 13, Application US/0938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/938,719
; APPLICATION NUMBER: US/09/938,719
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

DB: 10 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-938-703-13 (1-34)

QY 1 ATTAAGATAGTATCTGGGCTGCTCTCGCGCTGCTGTCATGTCATCTGCTACTC 60
|||||
Db 4 IleLysAspSerHisLeuGlyAlaGlyProAlaAlaCysHisGlyHisLeuLeu 23
|||||
QY 61 GGAATCTCTAAACTCTGCTCGGTGTCGAAA 93
|||||
Db 24 GlyAsnProLysAsnSerAlaSerValSerLys 34
|||||

RESULT 4

US-09-938-719-6

; Sequence 6, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/938/719
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/626,939
FILING DATE: 27-JULY-2000
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-938-719-6

Alignment Scores:
Pred. No.: 5,06e-15 Length: 215
Score: 164.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-938-719-6 (1-215)

QY 1 ATTAAGATAGTATCTGGGCTGCTCTCGCGCTGCTGTCATGTCATCTGCTACTC 60
|||||
Db 185 IleLysAspSerHisLeuGlyAlaGlyProAlaAlaCysHisGlyHisLeuLeu 204
|||||
QY 61 GGAATCTCTAAACTCTGCTCGGTGTCGAAA 93
|||||
Db 205 GlyAsnProLysAsnSerAlaSerValSerLys 215
|||||

US-09-938-703-3_COPY_792_885 (1-94) x US-09-938-703-13 (1-34)

QY 1 ATTAAGATAGTATCTGGGCTGCTCTCGCGCTGCTGTCATGTCATCTGCTACTC 60
|||||
Db 185 IleLysAspSerHisLeuGlyAlaGlyProAlaAlaCysHisGlyHisLeuLeu 204
|||||
QY 61 GGAATCTCTAAACTCTGCTCGGTGTCGAAA 93
|||||
Db 205 GlyAsnProLysAsnSerAlaSerValSerLys 215
|||||

RESULT 5

US-09-939-226-6

; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,226
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/626,939
FILING DATE: 2000-07-27
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-939-226-6

Alignment Scores:
Pred. No.: 5,06e-15 Length: 215
Score: 164.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-939-226-6 (1-215)

QY 1 ATTAAGATAGTATCTGGGCTGCTCTCGCGCTGCTGTCATGTCATCTGCTACTC 60
|||||
Db 185 IleLysAspSerHisLeuGlyAlaGlyProAlaAlaCysHisGlyHisLeuLeu 204
|||||
QY 61 GGAATCTCTAAACTCTGCTCGGTGTCGAAA 93
|||||
Db 205 GlyAsnProLysAsnSerAlaSerValSerLys 215
|||||

RESULT 6

US-09-938-703-6

; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR


```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-086-814-1

Alignment Scores:
Pred. No.: 2.68e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservatively: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 9 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-10-086-814-1 (1-352)
QY 2 TTAAGATAGTCATCTGGGGCTGCTCTCGCTGCTGTCATCTGCTACTCG 61
Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGTGTGCGAAT 94
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 12
US-09-734-221A-14
; Sequence 14, Application US/09734221A
; Publication No. US20030096221A1
; GENERAL INFORMATION:
; APPLICANT: LITTMAN, DAN R.
; DENG, HONGKUI
; ELLMEIER, WILFRIED
; LANDAU, NATHANIEL R.
; LIU, RONG
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH
; MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC
; USES THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/734,221A
; FILING DATE: 11-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/666,020
; FILING DATE: 19-JUN-1996
; APPLICATION NUMBER: US 08/227,319
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-004 N2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-734-221A-14

Alignment Scores:
Pred. No.: 2.68e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservatively: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 9 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-734-221A-14 (1-352)
QY 2 TTAAGATAGTCATCTGGGGCTGCTCTCGCTGCTGTCATCTGCTACTCG 61
Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGTGTGCGAAT 94
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226

RESULT 13
US-09-725-285-2
; Sequence 2, Application US/09725285
; Patent No. US20010000241A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruden, Steven, M.
; TITLE OF INVENTION: Antibodies to Human G-Protein Chemokine Receptor HDGNR10
; FILE REFERENCE: 1488.1150003
; CURRENT APPLICATION NUMBER: US/09/725,285
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 09/339,912
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/195,662
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 08/466,343
; PRIOR FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence: Genomic
; FEATURE:
; OTHER INFORMATION: Deduced Amino Acid Sequence
US-09-725-285-2

Alignment Scores:
Pred. No.: 2.68e-13 Length: 352
Score: 152.00 Matches: 31
Percent Similarity: 100.00% Conservatively: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.68% Indels: 0
DB: 10 Gaps: 0

US-09-938-703-3_COPY_792_885 (1-94) x US-09-725-285-2 (1-352)
QY 2 TTAAGATAGTCATCTGGGGCTGCTCTCGCTGCTGTCATCTGCTACTCG 61
Db 196 LeuLysIleValIleLeuGlyLeuValLeuProLeuLeuValMetValIleCysTyrSer 215
QY 62 GGAATCCTAAAACTCTGCTCGTGTGCGAAT 94
Db 216 GlyIleLeuLysThrLeuLeuArgCysArgAsn 226
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:19:16 ; Search time 6.42683 Seconds
(without alignments)
993.874 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 164 | 100.0 | 215 | 4 | 075303 |
| 2 | 115 | 70.1 | 33 | 4 | Q9P1F8 |
| 3 | 55 | 33.5 | 52 | 4 | Q9P1F5 |
| 4 | 55 | 33.5 | 178 | 4 | Q8TEJ8 |
| 5 | 55 | 33.5 | 336 | 16 | Q8VT73 |
| 6 | 54 | 32.9 | 276 | 5 | Q9W5C7 |
| 7 | 54 | 32.9 | 407 | 16 | Q8XNW7 |
| 8 | 54 | 32.9 | 727 | 2 | Q9AF65 |
| 9 | 53 | 32.3 | 312 | 16 | Q9LID3 |
| 10 | 53 | 32.3 | 499 | 16 | Q9A9Q4 |
| 11 | 52.5 | 32.0 | 478 | 10 | Q9LES2 |
| 12 | 52.5 | 32.0 | 478 | 10 | Q94B29 |
| 13 | 52 | 31.7 | 408 | 16 | Q97FK0 |
| 14 | 51 | 31.1 | 141 | 13 | Q9W619 |
| 15 | 51 | 31.1 | 187 | 3 | Q8WZT3 |
| 16 | 50 | 30.5 | 183 | 4 | Q9P098 |

| | | | | | |
|----|------|------|------|----|--------|
| 17 | 49.5 | 30.2 | 196 | 11 | Q8R3C9 |
| 18 | 49.5 | 30.2 | 756 | 11 | Q9CW46 |
| 19 | 49 | 29.9 | 331 | 10 | Q64731 |
| 20 | 49 | 29.9 | 515 | 12 | O40960 |
| 21 | 49 | 29.9 | 603 | 5 | Q9XVA3 |
| 22 | 48.5 | 29.6 | 500 | 6 | Q9TUJ5 |
| 23 | 48.5 | 29.6 | 698 | 5 | Q9W3F7 |
| 24 | 48 | 29.3 | 242 | 5 | O16761 |
| 25 | 48 | 29.3 | 325 | 11 | Q9QZP7 |
| 26 | 48 | 29.3 | 370 | 5 | Q95TX4 |
| 27 | 48 | 29.3 | 626 | 11 | Q9QKB9 |
| 28 | 48 | 29.3 | 773 | 16 | Q9L1F7 |
| 29 | 48 | 29.3 | 828 | 5 | Q9VKR7 |
| 30 | 48 | 29.3 | 989 | 2 | Q59736 |
| 31 | 48 | 29.3 | 1156 | 3 | Q8SQY1 |
| 32 | 48 | 29.3 | 1223 | 4 | Q8WXS8 |
| 33 | 48 | 29.3 | 1223 | 4 | Q8TE55 |
| 34 | 47.5 | 29.0 | 215 | 13 | Q9P093 |
| 35 | 47.5 | 29.0 | 677 | 10 | Q9FIM7 |
| 36 | 47.5 | 29.0 | 747 | 5 | Q9BIE6 |
| 37 | 47.5 | 29.0 | 971 | 5 | Q9VEC5 |
| 38 | 47 | 28.7 | 159 | 12 | Q997G6 |
| 39 | 47 | 28.7 | 170 | 12 | Q65868 |
| 40 | 47 | 28.7 | 217 | 5 | Q9VL41 |
| 41 | 47 | 28.7 | 246 | 10 | Q94E56 |
| 42 | 47 | 28.7 | 294 | 17 | Q9TD42 |
| 43 | 47 | 28.7 | 302 | 12 | O36364 |
| 44 | 47 | 28.7 | 307 | 5 | Q9W158 |
| 45 | 47 | 28.7 | 330 | 16 | Q8XRC0 |

ALIGNMENTS

RESULT 1

075303 PRELIMINARY; PRT; 215 AA.
AC 075303;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CC-chemokine receptor.
GN CCR-5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tse L., Ehrenberg P.K., Chang G., Michael N.L.;
RT *Genomic Organization and Functional Characterization of the Complete Transcription Unit for the Chemokine Receptor CCR-5, a Major Entry Co-Receptor for HIV-1.*
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009962; AAC23944.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 215 AA; 23946 MW; 3C9146C76BA416F7 CRC64;

Query Match 100.0%; Score 164; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.9e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31

Db 185 IKDHLGAGPAAACHGHLILGNPKNSASVSK 215

RESULT 2

Q9P1F8

| | | | | |
|-----------------------|---|----------------|----------------|-------------|
| ID | Q8TEU8 | PRELIMINARY; | PRT; | 178 AA. |
| AC | Q8TEU8; | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Created) | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Last sequence update) | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) | | | |
| DE | FLJ00197 protein (Fragment). | | | |
| DE | FLJ00197. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RC | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=SPLEEN; | | | |
| RA | Jikuya H., Takano J., Momura N., Kikuno R., Nagase T., Ohara O.; | | | |
| RT | "The nucleotide sequence of a long cDNA clone isolated from human | | | |
| RT | spleen."; | | | |
| RL | Submitted (JAN-2002) to the EMBL/GenBank/DDBT databases. | | | |
| RL | EMBL; AK074125; BAB84951.1; - | | | |
| FT | NON_TER 1 | | | |
| FT | SEQUENCE 178 AA; 18914 MW; E7BAB999932424CB CRC64; | | | |
| QY | 3 DSHLGAGPAAACHGHLILGNPKNS 26 | | | |
| | | | | |
| Db | 120 DSHLGGGPAATA-----GGPRTS 137 | | | |
| RESULT 5 | | | | |
| Q8YT73 | | | | |
| ID | Q8YT73 | PRELIMINARY; | PRT; | 336 AA. |
| AC | Q8YT73; | | | |
| DT | 01-MAR-2002 (TrEMBLrel. 20, Created) | | | |
| DT | 01-MAR-2002 (TrEMBLrel. 20, Last sequence update) | | | |
| DT | 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) | | | |
| DE | UDP-glucose 4-epimerase. | | | |
| DE | ALL2853. | | | |
| OS | Anabaena sp. (strain PCC 7120). | | | |
| OC | Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc. | | | |
| OX | NCBI_TaxID=103690; | | | |
| RN | [1] | | | |
| RC | SEQUENCE FROM N.A. | | | |
| RC | MEDLINE=21595285; PubMed=11759840; | | | |
| RA | Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S., | | | |
| RA | Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T., | | | |
| RA | Kisida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A., | | | |
| RA | Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yanada M., | | | |
| RA | Yasuda M., Tabata S.; | | | |
| RT | "Complete genomic sequence of the filamentous nitrogen-fixing | | | |
| RT | cyanobacterium Anabaena sp. strain PCC 7120."; | | | |
| RL | DNA Res. 8:205-213(2001). | | | |
| DR | EMBL: AP003591; BAB74552.1; - | | | |
| DR | InterPro; IPR001509; Epimerase_Dh. | | | |
| DR | InterPro; IPR000205; NAD-binding. | | | |
| DR | Pfam; PF01370; Epimerase; 1. | | | |
| DR | TIGRFAMS; TIGR011179; gale; 1. | | | |
| KW | Complete proteome. | | | |
| SQ | SEQUENCE 336 AA; 36733 MW; 5C4E7F6F6E855276 CRC64; | | | |
| Query Match | 33.5%; | Score 55; | DB 16; | Length 336; |
| Best Local Similarity | 35.5%; | Pred. No. 6.3; | | |
| Matches 11; | Conservative | 7; | Mismatches 13; | Indels 0; |
| QY | 1 IKDSHLGAGPAAACHGHLILGNPKNSASVSK 31 | | | |
| | | | | |
| Db | 31 VYDNCSTGSPQAVLHGELIIGDLKNSCLSQ 61 | | | |
| RESULT 6 | | | | |
| Q9W5C7 | | | | |

```

ID O9W5C7 PRELIMINARY; PRT; 276 AA.
AC O9W5C7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CG14633 protein.
GN CG14633.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE005419; AF45560.1; -
DR FlyBase; FBgn0029537; CG14633.
SQ SEQUENCE 276 AA; 28602 MW; 9F10B7DE62EA3A7 CRC64;

Query Match 32.9%; Score 54; DB 5; Length 276;
Best Local Similarity 42.4%; Pred. No. 7.3;
Matches 14; Conservative 1; Mismatches 2; Indels 16; Gaps 2;

OY 7 GAGPAAAC-----H-GHLLGNP 23
Db 167 GAGPAAACLVCCSPSSHHHGHVGHGATGHP 199
|||||
| |||||
| |||||

RESULT 7
O8XNW7 PRELIMINARY; PRT; 407 AA.
AC O8XNW7
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Probable exonuclease.
GN SBCD OR CPE0215.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yanashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003185; BAB79921.1; -
DR InterPro; IPR004843; M_ppestrase.
DR InterPro; IPR004844; S/T_phosphatase.
DR InterPro; IPR004593; Sbcd.
DR Pfam; PF00149; Metallophos; 1.
DR TIGRfams; TIGR00619; sbcd; 1.
KW Exonuclease; Complete proteome.
SQ SEQUENCE 407 AA; 46819 MW; 5E71E267DDC8B4A5 CRC64;

Query Match 32.9%; Score 54; DB 16; Length 407;
Best Local Similarity 38.5%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 6 LGAGPAAACHGHLILGNPKNSASVSK 31
: ||||| ||| :||| ||| :
Db 92 VAAGPLAREHGIIVGTPKSIIVEGE 117
: ||||| ||| :||| ||| :

RESULT 8
O9AF65 PRELIMINARY; PRT; 727 AA.
ID O9AF65
AC O9AF65;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cellulase Cella.
OC Clavibacter michiganensis subsp. sepe-donicus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Microbacteriaceae; Clavibacter.
OX NCBI_TaxID=31964;
RN [1]
RP SEQUENCE FROM N.A.
RA Laine M., Haapalainen M., Wahlroos T., Kankare K., Missinen R.,
RA Kassuwi S., Metzler M.C.;
RT "The cellulase encoded by the native plasmid of Clavibacter
RT michiganensis subsp. sepe-donicus plays a role in virulence and
RT contains an expansin-like domain."
RL Physiol. Mol. Plant Pathol. 0:0-0(2001).
DR EMBL; AY007311; AAK16222.1; -
DR HSP; P54583; LECE.
DR InterPro; IPR001919; Bac_cellose-bind.
DR InterPro; IPR001547; GH_5.
DR Pfam; PF00553; CBW_2; 1.
DR Pfam; PF00150; cellulase; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN 1.
SQ SEQUENCE 727 AA; 75811 MW; 96DF9B664873985D CRC64;

Query Match 32.9%; Score 54; DB 2; Length 727;
Best Local Similarity 52.4%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 9 GPAACGSHLLILGNPKNSASY 29
Db 568 GGAACGSHLLVTPGKGVKRV 588
|||||
| |||||
| |||||

RESULT 9

```

Q91LD3
ID Q91LD3 PRELIMINARY; PRT: 312 AA.
AC Q91LD3;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical protein SC01549.
GN SC01549 OR SC11.05C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seyger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT *Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).
RL Nature 417:141-147(2002).
DR EMBL; AL157953; CAB76068.1; -.
KW Hypothetical protein.
SQ SEQUENCE 312 AA; 32324 MW; A9E75F8C1DCE1C9F CRC64;

Query Match 32.3%; Score 53; DB 16; Length 312;
Best Local Similarity 76.9%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 GPAACHGHLHG 21
Db 179 GPAACHGHLHG 191
||||| ||| |||

RESULT 10
Q9A904
ID Q9A904 PRELIMINARY; PRT: 499 AA.
AC Q9A904;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein CC0920.
GN CC0920.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hart D.H.,
RA Klotz J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT *Complete genome sequence of Caulobacter crescentus.
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AC005768; AAK22904.1; -.
DR TIGR; CC0920; -.
DR InterPro; IPR001993; Mitoch. carrier.
DR PROSITE; PS00215; MITOCH. CARRIER; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 499 AA; 54933 MW; CCF42A7FB833CEAF CRC64;

Query Match 32.3%; Score 53; DB 16; Length 499;
Best Local Similarity 55.0%; Pred. No. 19;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 DSHLGAGPAACHGHLHG 22
Db 196 DARAGEGFARHCHGDLHGN 215
||| ||| ||| ||| |||

RESULT 11
Q9LES2
ID Q9LES2 PRELIMINARY; PRT: 478 AA.
AC Q9LES2;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 51.4 kDa protein (AT3G56860/T8M16_190) (UBP1 interacting
protein 2a).
GN T8M16_190 OR AT3G56860 OR UBA2A.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Wurmbech E., Drzonek H., Ansorge W., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Shinn P.,
RA Tracy S.E., Banh J., Bowser L., Carninci P., Chung M.K., Kamiya A.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jones T., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Pham P.K., Quach H.L.,
RA Sakano H., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT *Arabidopsis cDNA clones.
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT *Arabidopsis cDNA clones.
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Tang J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Deng C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT *Arabidopsis Full Length cDNA Clones.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Lambermon M.H.L., Fu Y., Wieczorek Kirk D.A., Dupasquier M.,
RA Filipowicz W., Lorkovic Z.J.;
RT *UBA1 and UBA2, two proteins that interact with UBPI, a

RT multifunctional effector of pre-mRNA maturation in plants.*;

RL Mol. Cell. Biol. 0:0-0(2002).
 DR EMBL: AL390921; CAC00749.1; -
 DR EMBL: AY037245; AAK59846.1; -
 DR EMBL: AF367280; AAK56269.1; -
 DR EMBL: AY091040; AAM13861.1; -
 DR EMBL: AY439404; CAD28672.1; -
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rtm; 2.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 478 AA; 51439 MW; 0CD3653669BE2DA4 CRC64;

Query Match 32.0%; Score 52.5; DB 10; Length 478;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDSHLCAGPAAACHGHLILGNP 23
 DB 346 RNDNNGYPPGG-HGHLMAGNP 366

RESULT 12

Q94B29 ID Q94B29 PRELIMINARY; PRT; 478 AA.
 AC Q94B29
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 51.4 kDa protein.
 GN TM6_190 OR AT355860.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

QY 2 KDSHLCAGPAAACHGHLILGNP 23
 DB 346 RNDNNGYPPGG-HGHLMAGNP 366

Query Match 32.0%; Score 52.5; DB 10; Length 478;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 2 KDSHLCAGPAAACHGHLILGNP 23
 DB 346 RNDNNGYPPGG-HGHLMAGNP 366

RESULT 13

Q97FK0 ID Q97FK0 PRELIMINARY; PRT; 408 AA.
 AC Q97FK0
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE DNA repair exonuclease.
 GN CAC2737.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RA "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum";
 RT J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007771; AAK80683.1; -
 DR InterPro: IPR004843; M-peptidase.
 DR InterPro: IPR004844; S/T_phosphatase.
 DR InterPro: IPR004593; Sbcd.
 DR Pfam: PF00149; Metallophos; 1.
 DR TIGRFAMs: TIGR00619; sbcd; 1.
 KW Exonuclease; Complete proteome.
 SQ SEQUENCE 408 AA; 46682 MW; CAEE3E17CFAE2B39 CRC64;

Query Match 31.7%; Score 52; DB 16; Length 408;
 Best Local Similarity 45.0%; Pred. No. 21;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 6 LGAGPAAACHGHLILGNP 25
 DB 92 VAAGPLARDHGIIMVGP 111

RESULT 14

Q9W619 ID Q9W619 PRELIMINARY; PRT; 141 AA.
 AC Q9W619
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE I22 protein.
 GN I22
 OS Cynops pyrrhogaster (Japanese common newt).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae; Cynops.
 OX NCBI_TaxID=8330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-REGENERATING LIMB BLASTOMA;
 RA Shimizu-Nishikawa K., Tazawa I., Uchiyama K., Yoshizato K.;
 RT "Expression of the helix-loop-helix type negative regulators of
 RT differentiation during limb regeneration in urodele and anuran";
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB019514; BAA76631.1; -
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR PROSITE: PS00036; HELIX_LOOP_HELIX; UNKNOWN_1.
 SQ SEQUENCE 141 AA; 15443 MW; C2012A212585C4A5 CRC64;

Query Match 31.1%; Score 51; DB 13; Length 141;
 Best Local Similarity 46.2%; Pred. No. 10;
 Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 3 DSHLGAGPAAACHGHLILGNPKNGAS 28
 ||| | | | | | : | |
 Db 83 DSHPGMGMAAGSGGHIQQRSFSASRS 108

RESULT 15

Q8WZT3 PRELIMINARY; PRT; 187 AA.
 AC Q8WZT3;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 21.3 kDa protein.
 GN B8L21.040.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL669989; CAD21090.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 187 AA; 21262 MW; 235332C807CC9F24 CRC64;

Query Match 31.1%; Score 51; DB 3; Length 187;
 Best Local Similarity 55.6%; Pred. No. 13;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 IKDSHLGAGPAAACHGHL 18
 | | | | : | | | | | : | :
 Db 49 IGTSHDSSPAAAHGDM 66

Search completed: June 3, 2003, 19:23:56
 Job time : 9.42683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:18:46 ; Search time 1.76423 Seconds
(without alignments)
728.798 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDSHGLGAGPAAACHGHLILGNPKNSAVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 56 | 34.1 | 240 | 1 YA02_SCHPO | Q9675 schizosacch |
| 2 | 51.5 | 31.4 | 681 | 1 HMC2_HELPF | Q9RQB4 helicobacte |
| 3 | 49.5 | 30.2 | 597 | 1 V7OK_OYMW | P20130 ononis yell |
| 4 | 48.5 | 29.6 | 222 | 1 NADD_XYLFA | Q9PBG5 xyliella fas |
| 5 | 48 | 29.3 | 358 | 1 INCL_ECOLI | Q52312 escherichia |
| 6 | 48 | 29.3 | 617 | 1 IF2P_AERPE | Q5YB33 aeropyrum p |
| 7 | 47.5 | 29.0 | 660 | 1 YFEG_ECO57 | Q8XZ33 escherichia |
| 8 | 47.5 | 29.0 | 660 | 1 YFEG_ECOLI | P77398 escherichia |
| 9 | 47.5 | 29.0 | 2319 | 1 YFEG_MOUSE | Q06194 mus musculu |
| 10 | 47.5 | 29.0 | 2359 | 1 CCAH_RAT | Q9EG60 rattus norv |
| 11 | 47 | 28.7 | 830 | 1 BGLS_BUTFI | P16084 butyrivibri |
| 12 | 46.5 | 28.4 | 516 | 1 CP23_HORVU | P52711 hordeum vul |
| 13 | 46.5 | 28.4 | 1059 | 1 Y379_HUMAN | Q15084 homo sapien |
| 14 | 46 | 28.0 | 211 | 1 CCMA_PARDE | P52218 paracoccus |
| 15 | 46 | 28.0 | 221 | 1 THIE_PASMU | P57930 pasteurella |
| 16 | 46 | 28.0 | 461 | 1 BENA_ACICA | P07769 acinetobact |
| 17 | 46 | 28.0 | 478 | 1 VP26_DROME | Q9K452 drosophila |
| 18 | 46 | 28.0 | 650 | 1 Y481_HUMAN | Q75069 homo sapien |
| 19 | 45.5 | 27.7 | 717 | 1 NFCL_MOUSE | Q88942 homo musculu |
| 20 | 45.5 | 27.7 | 816 | 1 QAIF_NEUCR | P11638 neurospora |
| 21 | 45.5 | 27.7 | 1011 | 1 HDA9_HUMAN | Q9UKV0 homo sapien |
| 22 | 45.5 | 27.7 | 1733 | 1 VNQA_PRYKA | P33485 pseudorabie |
| 23 | 45.5 | 27.7 | 2351 | 1 F48_HUMAN | P00451 homo sapien |
| 24 | 45 | 27.4 | 376 | 1 HMEV_DROME | P06602 drosophila |
| 25 | 45 | 27.4 | 815 | 1 CC53_YEAST | Q12018 saccharomyc |
| 26 | 45 | 27.4 | 1139 | 1 INAL_CABEL | Q03600 caenorhabdi |
| 27 | 44.5 | 27.1 | 660 | 1 YFPG_SALTI | Q8Z540 salmonella |
| 28 | 44.5 | 27.1 | 660 | 1 YFPG_SALTI | Q52325 salmonella |
| 29 | 44.5 | 27.1 | 661 | 1 TKL_TREPA | Q83571 treponema p |
| 30 | 44.5 | 27.1 | 670 | 1 TBUD_BURPI | Q01551 burkholderi |
| 31 | 44.5 | 27.1 | 841 | 1 NEK4_HUMAN | P51957 homo sapien |
| 32 | 44 | 26.8 | 188 | 1 RM32_HUMAN | Q9BYC8 homo sapien |
| 33 | 44 | 26.8 | 260 | 1 GTW1_CHICK | P20135 gallus gall |

34 44 26.8 420 1 RA52_MOUSE P43352 mus musculu
35 44 26.8 664 1 VP2_BPPH6 P11124 bacterioph
36 44 26.8 690 1 ATKB_SYNZ3 P73867 synechocyst
37 44 26.8 987 1 PMFG_CHLMU Q9P145 chlamydia m
38 43.5 26.5 309 1 YBGK_HAEIN P44298 haemophilus
39 43.5 26.5 786 1 AAC_ACTUT P29958 actinoplan
40 43.5 26.5 1127 1 MDW1_YEAST Q01846 saccharomyc
41 43 26.2 208 1 RISA_BUCAL P57212 buchnera ap
42 43 26.2 222 1 YA02_SCHPO Q10092 schizosacch
43 43 26.2 253 1 GS2_HUMAN P41247 homo sapien
44 43 26.2 286 1 RRPP_SYNV P21299 sonchus yel
45 43 26.2 353 1 XYNA_PENCH P29417 penicillium

ALIGNMENTS

RESULT 1
YA02_SCHPO STANDARD; PRT; 240 AA.
AC Q09675;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C5H10.02c in chromosome I.
GN SPAC5H10.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., Del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe R., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
The genome sequence of Schizosaccharomyces pombe.;
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: BELONGS TO THE YMR322C/YOR391C/YPL280W (YEAST),
CC SPAC1F7.06/SPAC5H10.02C/SPAC1D3.13 (S.POMBE FAMILY).
CC -!- SIMILARITY: SOME, TO E.COLI YEDU.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

QY 4 SHLGAGPAAACHGHLIL-----GNPKNSAS 28
||| : ||| | ||| : |||


```

GN  INFIB OR APE2374.
OS  Aeropyrum pernix.
OC  Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OX  Desulfurococcaceae; Aeropyrum.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-K1;
RX  MEDLINE-99310339; PubMed-10382966;
RA  Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA  Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA  Hosoyama M., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA  Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA  Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA  Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT  "Complete genome sequence of an aerobic hyper-thermophilic
RT  Crenarchaeon, Aeropyrum pernix K1.";
RL  DNA Res. 6:83-101(1999).
CC  -1- FUNCTION: FUNCTION IN GENERAL TRANSLATION INITIATION BY PROMOTING
CC  THE BINDING OF THE FORMYL METHIONINE-TRNA TO RIBOSOMES. SEEMS TO
CC  FUNCTION ALONG WITH EIF-2 (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AP000064; BAA81387.1; -
DR  InterPro: IPR004161; EFTU_D2
DR  InterPro: IPR000795; EF_GTPbind.
DR  InterPro: IPR000178; IF2.
DR  InterPro: IPR005225; Small_GTP.
DR  InterPro: IPR004544; TIF_aif-2.
DR  Pfam: PF00009; GTP_EFTU; 1.
DR  Pfam: PF03144; GTP_EFTU_D2; 1.
DR  PRINTS; PR00315; ELONGATNCT.
DR  TIGRFAMs; TIGR00231; small_gtp; 1.
DR  TIGRFAMs; TIGR00491; aif-2; 1.
DR  PROSITE; PS01176; IF2; FALSE NEG.
KW  Initiation factor; Protein biosynthesis; GTP-binding;
KW  Complete proteome.
FT  NP_BIND 23 30 GTP (BY SIMILARITY).
FT  NP_BIND 87 91 GTP (BY SIMILARITY).
FT  NP_BIND 141 144 GTP (BY SIMILARITY).
SQ  SEQUENCE 617 AA; 68520 MW; 493DC149C6028EB3 CRC64;

Query Match 29.3%; Score 48; DB 1; Length 617;
Best Local Similarity 27.6%; Pred. No. 20;
Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 IKDSHIGAGPAAACHGHLILGNPKNSASV 29
D531 LEEARLGAARVAVSQIRLLIGRANEEDI 559

RESULT 7
YFBG_EC057
ID YFBG_EC057 STANDARD; PRT; 660 AA.
AC Q8XDZ3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yfbg.
GN YFBG OR Z3513 OR EC33143.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=83334;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE-21074935; PubMed-11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE-21156231; PubMed-11238796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- SIMILARITY: BELONGS TO THE FMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF005458; AAG57386.1; -
DR EMBL; AP002561; BAB36566.1; -
DR InterPro: IPR002376; formyl_transf.
DR InterPro: IPR004187; Formyl_transf_C.
DR Pfam: PF00551; formyl_transf; 1.
DR Pfam: PF02911; formyl_transf_C; 1.
KW Hypothetical protein; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 660 AA; 74209 MW; 9D45B3AF60CF6131 CRC64;

Query Match 29.0%; Score 47.5; DB 1; Length 660;
Best Local Similarity 42.1%; Pred. No. 25;
Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHL-LGNPKNSASVSK 31
D538 CDGEIINIGNPENEASIE 576

RESULT 8
YFBG_EC01
ID YFBG_EC01 STANDARD; PRT; 660 AA.
AC P77398;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yfbg.
GN YFBG OR B2255.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";

```



```

FT CARBOHYD 380 380 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 516 AA; 55913 MW; D41A1C56CF9D355 CRC64;

Query Match 28.4%; Score 46.5; DB 1; Length 516;
Best Local Similarity 40.6%; Pred. No. 28;
Matches 13; Conservative 3; Mismatches 7; Indels 9; Gaps 1;

QY 6 LGAGPAAACG-----HLLGNPKNSAS 28
      |||||
DB 19 LGAGPAAAKARETRQGDYLNRLRGSPSSRAS 50

RESULT 13
Y379_HUMAN
ID Y379_HUMAN STANDARD; PRT; 1059 AA.
AC O15084;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA0379 (Fragment).
GN KIAA0379
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9734984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
CC
CC -!- SIMILARITY: CONTAINS AT LEAST 27 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AB002377; BAA20833.2; -
CC HSSP; P80144; 2MYO.
CC InterPro; IPR002110; ANK.
CC Pfam; PF00023; ank; 28.
CC SMART; SM00248; ANK; 26.
CC PROSITE; PS50088; ANK_REPEAT; 24.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
KW NON_TER 1
FT REPEAT 46 FT 75 ANK 1.
FT REPEAT 79 108 ANK 2.
FT REPEAT 112 141 ANK 3.
FT REPEAT 145 174 ANK 4.
FT REPEAT 178 207 ANK 5.
FT REPEAT 211 240 ANK 6.
FT REPEAT 244 273 ANK 7.
FT REPEAT 277 307 ANK 8.
FT REPEAT 311 340 ANK 9.
FT REPEAT 344 373 ANK 10.
FT REPEAT 377 406 ANK 11.
FT REPEAT 410 439 ANK 12.
FT REPEAT 443 472 ANK 13.
FT REPEAT 476 506 ANK 14.
FT REPEAT 510 540 ANK 15.
FT REPEAT 555 584 ANK 16.
FT REPEAT 588 617 ANK 17.
FT REPEAT 622 651 ANK 18.
FT REPEAT 658 687 ANK 19.

```

```

Db      163 AHIGGGAATVATHIDGLPE 183

RESULT 15
THIE_PASMO
ID THIE_PASMO STANDARD; PRT; 221 AA.
AC P57930;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thiamine-phosphate pyrophosphorylase (EC 2.5.1.3) (TMP
DE pyrophosphorylase) (TMP-PPase) (Thiamine-phosphate synthase).
GN THIE OR PM1260.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RF "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: CONDENSES 4-METHYL-5-(BETA-HYDROXYETHYL)-THIAZOLE
CC MONOPHOSPHATE (THZ-P) AND 4-AMINO-5-HYDROXYMETHYL PYRIMIDINE
CC PYROPHOSPHATE (HMP-PP) TO FORM THIAMINE MONOPHOSPHATE (TMP) (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2-methyl-4-amino-5-hydroxymethylpyrimidine
CC diphosphate + 4-4-methyl-5-(2-phosphonooxyethyl)-thiazole =
CC diphosphate + thiamine monophosphate.
CC -1- PATHWAY: Thiamine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE TMP-PPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006165; AAK03344.1; -.
DR HSSP; P39594; 2TPS.
DR InterPro; IPR003733; TMP_synthase.
DR Pfam; PF02581; TMP-TENI; 1.
DR TIGRams; TIGR00693; thie; 1.
KW Thiamine biosynthesis; transferase; Complete proteome.
SQ SEQUENCE 221 AA; 23736 MW; 2207E7CF5E62134 CRC64;

Query Match 28.0%; Score 46; DB 1; Length 221;
Best Local Similarity 35.7%; Pred. No. 14;
Matches 10; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 5 HLG-----AGPAACHGHLIGNPKNS 26
|:| | | | | | | | | |
Db 97 HVGQTDMAVDVAALCHSHCFGTGTSVT 124

Search completed: June 3, 2003, 19:22:57
Job time : 2.76423 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:24:01 ; Search time 7.43496 Seconds
(without alignments)
422.051 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215

Perfect score: 164

Sequence: 1 IKDHLGAGPAAACHGILLGNPKNSASVSK 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 164 | 100.0 | 34 | 10 | US-09-938-719-13 |
| 2 | 164 | 100.0 | 34 | 10 | US-09-939-226-13 |
| 3 | 164 | 100.0 | 34 | 10 | US-09-938-703-13 |
| 4 | 164 | 100.0 | 215 | 10 | US-09-938-719-6 |
| 5 | 164 | 100.0 | 215 | 10 | US-09-939-226-6 |
| 6 | 164 | 100.0 | 215 | 10 | US-09-938-703-6 |
| 7 | 55 | 33.5 | 84 | 10 | US-09-864-761-43094 |
| 8 | 48 | 29.3 | 297 | 10 | US-09-938-330-4 |
| 9 | 48 | 29.3 | 451 | 10 | US-09-938-330-2 |
| 10 | 48 | 29.3 | 486 | 10 | US-09-938-330-6 |
| 11 | 48 | 29.3 | 1213 | 10 | US-09-938-330-14 |
| 12 | 48 | 29.3 | 1216 | 10 | US-09-938-330-12 |
| 13 | 48 | 29.3 | 1219 | 10 | US-09-938-330-10 |
| 14 | 48 | 29.3 | 1222 | 10 | US-09-938-330-8 |
| 15 | 48 | 29.3 | 1223 | 9 | US-10-205-368-4 |
| 16 | 48 | 29.3 | 1223 | 12 | US-10-014-070-2 |
| 17 | 48 | 29.3 | 1232 | 10 | US-09-938-330-18 |
| 18 | 48 | 29.3 | 1233 | 9 | US-10-205-368-2 |
| 19 | 48 | 29.3 | 1235 | 10 | US-09-938-330-16 |

Sequence 22, Appl
Sequence 20, Appl
Sequence 5, Appl
Sequence 62, Appl
GENERAL INFORMA
Sequence 1726, Ap
Sequence 212, Ap
Sequence 3578, Ap
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 1655, Ap
Sequence 95, Appl
Sequence 10, Appl
Sequence 99, Appl
Sequence 11, Appl
Sequence 93, Appl

Sequence 22, Appl
Sequence 20, Appl
Sequence 5, Appl
Sequence 62, Appl
GENERAL INFORMA
Sequence 1726, Ap
Sequence 212, Ap
Sequence 3578, Ap
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 1655, Ap
Sequence 95, Appl
Sequence 10, Appl
Sequence 99, Appl
Sequence 11, Appl
Sequence 93, Appl

Sequence 22, Appl
Sequence 20, Appl
Sequence 5, Appl
Sequence 62, Appl
GENERAL INFORMA
Sequence 1726, Ap
Sequence 212, Ap
Sequence 3578, Ap
Sequence 2, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 1655, Ap
Sequence 95, Appl
Sequence 10, Appl
Sequence 99, Appl
Sequence 11, Appl
Sequence 93, Appl

ALIGNMENTS

RESULT 1
US-09-938-719-13
; Sequence 13, Application US/09938719
; Patent No. US20020106742A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/938,719
; APPLICATION NUMBER: US/09/938,719
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 27-JULY-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-938-719-13

Query Match 100.0%; Score 164; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGHLILGNPKNSASVSK 34

RESULT 2

US-09-939-226-13

; Sequence 13, Application US/0939226

; Patent No. US20020110805A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARMENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/939,226

FILING DATE: 24-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/626,939

FILING DATE: 2000-07-27

ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E

REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-939-226-13

Query Match 100.0%; Score 164; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGHLILGNPKNSASVSK 34

RESULT 3

US-09-938-703-13

; Sequence 13, Application US/0938703

; Patent No. US20020110870A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARMENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/938,703

FILING DATE: 24-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/626,939

FILING DATE: 2000-07-27

ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E

REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-09-938-703-13

Query Match 100.0%; Score 164; DB 10; Length 34;
Best Local Similarity 100.0%; Pred. No. 5.9e-17;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDHLGAGPAAACHGHLILGNPKNSASVSK 31
DB 4 IKDHLGAGPAAACHGHLILGNPKNSASVSK 34

RESULT 4

US-09-938-719-6

; Sequence 6, Application US/0938719

; Patent No. US20020106742A1

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; PARMENTIER, MARC

; VASSART, GILBERT

; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: U.S.A.

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/938,719

FILING DATE: 24-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/626,939
FILING DATE: 27-JULY-2000
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: <Unknown>
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-719-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDSLHAGPAAACHGHLGNPKNSASVSK 31
Db 185 IKDSLHAGPAAACHGHLGNPKNSASVSK 215

RESULT 5
US-09-939-226-6
; Sequence 6, Application US/09939226
; Patent No. US20020110805A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/939,226
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-939-226-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDSLHAGPAAACHGHLGNPKNSASVSK 31

Db 185 IKDSLHAGPAAACHGHLGNPKNSASVSK 215

RESULT 6
US-09-938-703-6
; Sequence 6, Application US/09938703
; Patent No. US20020110870A1
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; PARMENTIER, MARC
; VASSART, GILBERT
; LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/938,703
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/626,939
; FILING DATE: 2000-07-27
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-938-703-6

Query Match 100.0%; Score 164; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKDSLHAGPAAACHGHLGNPKNSASVSK 31
Db 185 IKDSLHAGPAAACHGHLGNPKNSASVSK 215

RESULT 7
US-09-864-761-43094
; Sequence 43094, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 43094
/ LENGTH: 84
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AL096816.12
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 2.4
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.93
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
/ OTHER INFORMATION: EST_HUMAN HIT: A1696698.1, EVALU6 7.00e-28
US-09-864-761-43094

Query Match      33.5%; Score 55; DB 10; Length 84;
Best Local Similarity 50.0%; Pred. No. 0.98;
Matches 12; Conservative 1; Mismatches 5; Indels 6; Gaps 1;

QY      3 DSHLGAGPAAACHGHLILGNPKNS 26
        ||||| |||||
DB      48 DSHLGGGPAATA-----GGPRTS 65

RESULT 8
US-09-938-330-4
/ Sequence 4, Application US/09938330
/ Patent No. US20020115838A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Hilbun, Erin
/ APPLICANT: Scoville, John
/ APPLICANT: Fiddle, Carl Johan
/ APPLICANT: Hu, Yi
/ APPLICANT: Turner, C. Alexander Jr.
/ TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encodi
/ FILE REFERENCE: LEX-0237-USA
/ CURRENT APPLICATION NUMBER: US/09/938,330
/ CURRENT FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 2001-08-22
```

```
/ PRIOR APPLICATION NUMBER: US 60/227,104
/ PRIOR FILING DATE: 2000-08-22
/ PRIOR APPLICATION NUMBER: US 60/233,796
/ PRIOR FILING DATE: 2000-09-19
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 297
/ TYPE: PRT
/ ORGANISM: homo sapiens
US-09-938-330-4

Query Match      29.3%; Score 48; DB 10; Length 297;
Best Local Similarity 45.0%; Pred. No. 43;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
        ||: ||||| | :|
DB      54 SHVSGPAAASAGSMVVDTP 73

RESULT 9
US-09-938-330-2
/ Sequence 2, Application US/09938330
/ Patent No. US20020115838A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Hilbun, Erin
/ APPLICANT: Scoville, John
/ APPLICANT: Fiddle, Carl Johan
/ APPLICANT: Hu, Yi
/ APPLICANT: Turner, C. Alexander Jr.
/ TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Enco
/ FILE REFERENCE: LEX-0237-USA
/ CURRENT APPLICATION NUMBER: US/09/938,330
/ CURRENT FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: US 60/227,104
/ PRIOR FILING DATE: 2000-08-22
/ PRIOR APPLICATION NUMBER: US 60/233,796
/ PRIOR FILING DATE: 2000-09-19
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 451
/ TYPE: PRT
/ ORGANISM: homo sapiens
US-09-938-330-2

Query Match      29.3%; Score 48; DB 10; Length 451;
Best Local Similarity 45.0%; Pred. No. 70;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
        ||: ||||| | :|
DB      54 SHVSGPAAASAGSMVVDTP 73

RESULT 10
US-09-938-330-6
/ Sequence 6, Application US/09938330
/ Patent No. US20020115838A1
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Hilbun, Erin
/ APPLICANT: Scoville, John
/ APPLICANT: Fiddle, Carl Johan
/ APPLICANT: Hu, Yi
/ APPLICANT: Turner, C. Alexander Jr.
/ TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Enco
/ FILE REFERENCE: LEX-0237-USA
/ CURRENT APPLICATION NUMBER: US/09/938,330
/ CURRENT FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: US 60/227,104
```

```
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 486
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-6

Query Match      29.3%  Score 48; DB 10; Length 486;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
||: :||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 11
US-09-938-330-14
; Sequence 14, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-14

Query Match      29.3%  Score 48; DB 10; Length 1213;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
||: :||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 12
US-09-938-330-12
; Sequence 12, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22

; PRIOR FILING DATE: 2000-08-22

; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-12

Query Match      29.3%  Score 48; DB 10; Length 1213;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
||: :||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 13
US-09-938-330-10
; Sequence 10, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-10

Query Match      29.3%  Score 48; DB 10; Length 1219;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
||: :||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73

RESULT 14
US-09-938-330-8
; Sequence 8, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1el Human Proteases and Polynucleotides Enco
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-8

Query Match      29.3%  Score 48; DB 10; Length 1219;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 SHLGAGPAAACHGHLILGNP 23
||: :||||| |:: |
Db      54 SHVYSGPAAASAGSMVVDTP 73
```


; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-8

Query Match 29.3%; Score 48; DB 10; Length 1222;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHLGAGPAAACHGHLILGNP 23
||: ||||| |:: |
Db 54 SHVSGPAAASAGSMVVDTP 73

RESULT 15

US-10-205-368-4
; Sequence 4, Application US/10205368
; Publication NO. US20030092621A1
; GENERAL INFORMATION:
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: RACIE, LISA A.
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0094-00000
; CURRENT APPLICATION NUMBER: US/10/205,368
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/308,386
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/315,887
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-368-4

Query Match 29.3%; Score 48; DB 9; Length 1223;
Best Local Similarity 45.0%; Pred. No. 2.2e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHLGAGPAAACHGHLILGNP 23
||: ||||| |:: |
Db 54 SHVSGPAAASAGSMVVDTP 73

Search completed: June 3, 2003, 19:36:00
Job time : 8.43496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 19:20:46 ; Search time 3.40244 Seconds
(without alignments)
875.891 Million cell updates/sec

Title: US-09-938-703-6_COPY_185_215
Perfect score: 164
Sequence: 1 IKDHLGAGPAAACHGHLGNPKNSASVSK 31

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR73:*
1: piri:*
2: piri:*
3: piri:*
4: piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|--------|--------------------|
| 1 | 56 | 34.1 | 240 | S55480 | hypothetical prote |
| 2 | 55 | 33.5 | 336 | AF2162 | UDP-glucose 4-epim |
| 3 | 53 | 32.3 | 499 | D87363 | conserved hypothet |
| 4 | 52.5 | 32.0 | 478 | T51274 | hypothetical prote |
| 5 | 52 | 31.7 | 408 | H97236 | DNA repair exonuc |
| 6 | 49.5 | 30.2 | 597 | J00107 | hypothetical 66K p |
| 7 | 49 | 29.9 | 331 | T00582 | probable mitochond |
| 8 | 48.5 | 29.6 | 222 | H82590 | conserved hypothet |
| 9 | 48.5 | 29.6 | 500 | J02059 | cytochrome P450 2D |
| 10 | 48 | 29.3 | 242 | T32122 | hypothetical prote |
| 11 | 48 | 29.3 | 358 | T08477 | inclusion membrani |
| 12 | 48 | 29.3 | 617 | C72466 | probable translati |
| 13 | 47.5 | 29.0 | 660 | G31021 | probable transform |
| 14 | 47.5 | 29.0 | 660 | F85865 | probable transform |
| 15 | 47.5 | 29.0 | 660 | E64996 | hypothetical prote |
| 16 | 47.5 | 29.0 | 2319 | A47004 | coagulation factor |
| 17 | 47 | 28.7 | 294 | G72706 | hypothetical prote |
| 18 | 47 | 28.7 | 302 | T03109 | probable membrane |
| 19 | 47 | 28.7 | 830 | A44768 | beta-glucosidase (|
| 20 | 46.5 | 28.4 | 516 | S44191 | carboxypeptidase D |
| 21 | 46 | 28.0 | 388 | T36766 | probable two-compo |
| 22 | 46 | 28.0 | 461 | S23477 | probable benzoate |
| 23 | 46 | 28.0 | 469 | T40339 | ferredoxin-NADP re |
| 24 | 46 | 28.0 | 478 | T13683 | embryogenesis prot |
| 25 | 46 | 28.0 | 884 | T18649 | hypothetical prote |
| 26 | 45.5 | 27.7 | 718 | JC5805 | transcription fact |
| 27 | 45.5 | 27.7 | 745 | B84673 | hypothetical prote |
| 28 | 45.5 | 27.7 | 816 | F31277 | regulatory protein |
| 29 | 45.5 | 27.7 | 1733 | B45344 | probable nuclear a |

30 45.5 27.7 1958 2 B40505
31 45.5 27.7 2351 1 EZHU
32 45 27.4 106 2 I48862
33 45 27.4 182 2 T06978
34 45 27.4 199 2 T44524
35 45 27.4 322 2 AE3093
36 45 27.4 329 2 D96030
37 45 27.4 340 2 E98193
38 45 27.4 376 2 A26066
39 45 27.4 390 2 T51197
40 45 27.4 437 2 A36372
41 45 27.4 484 2 T00158
42 45 27.4 511 2 H89758
43 45 27.4 638 2 T26490
44 45 27.4 675 2 I39065
45 45 27.4 815 2 S67675

ALIGNMENTS

RESULT 1

S55480
hypothetical protein SPAC5H10.02c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 10-Dec-1999
C:Accession: S55480; T38966
R:Connor, R.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55479
A:Accession: S55480
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-240 <CON>
A:Cross-references: EMBL:Z49811; NID:g854599; PIDN:CAA89952.1; PID:g854601
R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsby, S.V.
submitted to the EMBL Data Library, May 1995
A:Reference number: Z21821
A:Accession: T38966
A>Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-240 <CO2>
A:Cross-references: EMBL:Z49811; PIDN:CAA89952.1; GSPDB:GN00066; SPDB:SPAC5H10.02c
A:Experimental source: strain 972h-; cosmid c5H10
C:Genetics:
A:Gene: SPDB:SPAC5H10.02c
A:Map position: 1
C:Superfamily: conserved hypothetical protein YMR322c

Query Match 34.1%; Score 56; DB 2; Length 240;
Best Local Similarity 57.9%; Pred.No. 2.7;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 11 AAACHGHLGNPKNSASV 29
:||||:|||||
Db 138 SAVCHGPVLLANVKNPQSV 156

RESULT 2

UDP-glucose 4-epimerase [Imported] - Nostoc sp. (strain PCC 7120)
AF2162
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AF2162
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irig
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabat
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2162
A>Status: preliminary
A:Molecule type: DNA

RESULT 7
T00582
probable mitochondrial carrier protein [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T27E13.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00582; C84705
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.

Submitted to the EMBL data library, May 1996
A: Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.
A: Reference number: Z14178
A: Accession: T00582
A: Status: translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-331 <R07>
A: Cross-references: EMBL:AC004165; NID:g3150396; PID:g3150404
A: Experimental source: cultivar Columbia
A: Lili, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Penito, M.I.; Town, C.D.; Fut

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: R84420; MUID:20083487; PMID:10617197

A:Accession: C84705

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <STO>

A:Cross-references: GB:AE002093; NID:g3150404; PIDN:AAC16956.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g30160; T27E13.10

A:Map position: 2

A:Introns: 263/3

C:Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology

C:Keywords: mitochondrion

Query Match 29.9%; Score 49; DB 2; Length 331;

Best Local Similarity 45.2%; Pred. No. 35;

Matches 14; Conservative 2; Mismatches 7; Indels 8; Gaps 1;

QY 6 LGAGPAAACH-----GHLGPNKNSAS 28

DB 104 LGAGPAHAYFSEVSKFLSGGNPNNSAA 134

RESULT 8

H82590

C:Species: conserved hypothetical protein XF2179 [Imported] - *Xylella fastidiosa* (strain 9a5c)

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 01-Mar-2002

C:Accession: H82590

R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: H82590

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-222 <SIM>

A:Cross-references: GB:AE004031; GB:AE003849; NID:g9107309; PIDN:AAF84978.1; GSPDB:GN001

A:Experimental source: Strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2179

C:Superfamily: *Bacillus subtilis* conserved hypothetical protein yqeJ

Query Match 29.6%; Score 48.5; DB 2; Length 222;

Best Local Similarity 61.1%; Pred. No. 28;

Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 15 HGHLLGNPKNSAVSK 31

DB 170 HGHLLGNPKNSAVSK 187

RESULT 9

JE0259

C:Species: cytochrome P450 2D24 - rabbit

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: *Oryctolagus cuniculus* (domestic rabbit)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000

C:Accession: JE0259

R:Yamamoto, Y.; Ishizuka, M.; Takada, A.; Fujita, S.

J. Biochem. 124, 503-508, 1998

A:Title: Cloning, tissue distribution, and functional expression of two novel rabbit

A:Reference number: JE0258; MUID:98391821; PMID:9722658

A:Accession: JE0259

A:Molecule type: mRNA

A:Residues: 1-500 <YAM>

A:Cross-references: DDBJ:AB008785

C:Comment: This protein shows high drug metabolizing activity.

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; met

F:9-25/Domain: transmembrane #status predicted <TM1>

F:305-468/Domain: cytochrome P450 homology <CYP>

F:310-326/Domain: transmembrane #status predicted <TM2>

F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 29.6%; Score 48.5; DB 1; Length 500;

Best Local Similarity 55.6%; Pred. No. 61;

Matches 10; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 5 HLGAGPAA-----ACHGH 17

DB 112 HLGFQAGQVIMACYGH 129

RESULT 10

T32122

C:Species: hypothetical protein F59E11.5 - *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32122

R:Bradshaw, H.

A:Description: The sequence of *C. elegans* cosmid F59E11.

A:Reference number: Z21124

A:Accession: T32122

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-242 <BRA>

A:Cross-references: EMBL:AF016685; PIDN:AAB66219.1; GSPDB:GN00023; CESP:F59E11.5

A:Experimental source: strain Bristol N2; clone F59E11

C:Genetics:

A:Gene: CESP:F59E11.5

A:Map position: 5

A:Introns: 87/3; 123/3; 180/3; 217/3

Query Match 29.3%; Score 48; DB 2; Length 242;

Best Local Similarity 39.3%; Pred. No. 36;

Matches 11; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 2 KDHLGAGPAAACHGHLGPNKNSAV 29

DB 88 KDHLGAGPAAACHGHLGPNKNSAV 115

RESULT 11

T08477

C:Species: inclusion membrane protein C1 - *Enterobacter aerogenes* plasmid R751

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: T08477

R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.

Plasmid 36, 95-111, 1996

A:Title: Conservation of the genetic switch between replication and transfer genes o

A:Reference number: Z16434; MUID:97118926; PMID:8954881

A:Accession: T08477

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-358 <THO>

Wed Jun 11 09:06:35 2003

A:Cross-references: EMBL:U67194; NID:g1572520; PIDN:ARC64421.1; PID:g1572526

C:Genetics:

A:Gene: incC1

A:Genome: plasmid R751

C:Superfamily: incC protein

Query Match 29.3%; Score 48; DB 2; Length 358;
 Best Local Similarity 58.8%; Pred. No. 52;
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 IKDSHLGAGPAAACHGH 17

DB 14 IPDGHQAGDRAADHRH 30

RESULT 12

C72466 Probable translation initiation factor aIF-2 APE2374 [similarity] - Aeropyrum pernix (st

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Species: Aeropyrum pernix

C:Accession: C72466

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72466

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-617 <RAW>

A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81387.1; PID:g5106076

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2374

C:Superfamily: translation initiation factor IF-2; translation elongation factor Tu hom

Query Match 29.3%; Score 48; DB 2; Length 617;
 Best Local Similarity 27.6%; Pred. No. 88;
 Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 IKDSHLGAGPAAACHGHLGNPKNSASV 29

DB 531 LEEARLGA AVAVSIQGRILIGHANEGDI 559

RESULT 13

G91021 Probable transformylase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: G91021

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: G91021

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-660 <RAY>

A:Cross-references: PIDN:BA836566.1; PID:g13362613; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 050952

C:Genetics:

A:Gene: ECs3143

Query Match 29.0%; Score 47.5; DB 2; Length 660;
 Best Local Similarity 42.1%; Pred. No. 1.1e+02;
 Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSK 31

DB 558 CDGEIINIGNPENEASIEE 576

RESULT 14

F85865

Probable transformylase Z3513 [imported] - Escherichia coli (strain O157:H7, substra

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: F85865

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F85865

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-660 <STO>

A:Cross-references: GB:AE005174; NID:g12516601; PIDN:AA657386.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z3513

Query Match 29.0%; Score 47.5; DB 2; Length 660;
 Best Local Similarity 42.1%; Pred. No. 1.1e+02;
 Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSK 31

DB 558 CDGEIINIGNPENEASIEE 576

RESULT 15

E64996

Hypothetical protein b2255 - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: E64996

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: E64996

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-660 <BLAT>

A:Cross-references: GB:AE000315; GB:U00096; NID:g1788582; PIDN:AACT5315.1; PID:g17885

A:Experimental source: strain K-12, substrain MG1655

Query Match 29.0%; Score 47.5; DB 2; Length 660;
 Best Local Similarity 42.1%; Pred. No. 1.1e+02;
 Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 14 CHGHLL-LGNPKNSASVSK 31

DB 558 CDGEIINIGNPENEASIEE 576

Search completed: June 3, 2003, 19:24:29

Job time : 5.40244 secs